

С. Э. НОХРИН

МАТЕМАТИКА ДЛЯ ЭКОНОМИСТОВ

Конспект лекций



Министерство образования и науки Российской Федерации
Уральский федеральный университет
имени первого Президента России Б. Н. Ельцина

МАТЕМАТИКА ДЛЯ ЭКОНОМИСТОВ

Курс лекций

*Рекомендовано методическим советом УрФУ
для студентов, обучающихся по направлениям
080100.62 — Экономика, 080200.62 — Менеджмент,
230700.62 — Прикладная информатика,
080500.62 — Бизнес-информатика*

Екатеринбург
Издательство Уральского университета
2014

УДК 51-7:330(075.8)

ББК 22.1я73+65я73

Н85

Составитель — С. Э. Нохрин

Рецензенты:

д-р физ.-мат. наук *А. В. Осипов* (Институт математики и механики им. Н. Н. Красовского);

канд. физ.-мат. наук, доц. *С. А. Ануфриенко* (Специализированный учебно-научный центр УрФУ)

Научный редактор — канд. физ.-мат. наук *О. Я. Шевалдина*

Нохрин, С. Э.

Н85 Математика для экономистов : курс лекций / сост. С. Э. Нохрин. — Екатеринбург : Изд-во Урал. ун-та, 2014. — 120 с.

ISBN 978-5-7996-1251-1

Пособие состоит из лекций, читаемых в процессе курса «Математика для экономистов» для студентов первого курса технических специальностей. Пособие может быть использовано для самостоятельного изучения предмета и для ликвидации пробелов в курсе алгебры средней школы.

Подготовлено кафедрой «Моделирование управляемых систем»

Библиогр.: 5 назв. Табл. 1. Рис. 17.

УДК 51-7:330(075.8)

ББК 22.1я73+65я73

ISBN 978-5-7996-1251-1

© Уральский федеральный
университет, 2014

ЛЕКЦИЯ I. ОСНОВЫ ТЕОРИИ МНОЖЕСТВ

1. Понятие множества

Множество — есть понятие неопределяемое. Его можно трактовать как некоторый набор объектов. Эти объекты называются *элементами* множества. Множество считается заданным, если про каждый объект можно сказать, является он элементом данного множества или не является. Так, можно говорить о множестве студентов УрФУ, о множестве треугольников, лежащих в данной плоскости или о множестве решений какого-либо уравнения. В последнем случае заметим, что совершенно неважно, умеем мы решать указанное уравнение или нет; множество его корней определено однозначно и может являться предметом математического исследования.

Поведение множеств описывается с помощью аксиом теории множеств. Общепринятая в математике система аксиом теории множеств — *аксиомы Цермело Френкеля*. Эта система обозначается ZFC . К этой системе аксиом, как правило, добавляется так называемая *аксиома выбора*. В настоящем курсе мы не будем вникать в тонкости аксиоматики; желающие могут подробно ознакомиться с нею, изучив соответствующую литературу, например, «Справочную книгу по математической логике» (часть 2 под редакцией Дж. Барвайса).

Для обозначения множеств применяются заглавные латинские буквы (с индексами или без них), например A , B_{12} или T_n . Аналогично, элементы множеств, как правило, будут обозначаться строчными латинскими буквами (возможно, также с индексами). Следует отметить, что сами множества также могут являться элементами каких-то множеств, например, любая прямая — это некоторое множество точек, поэтому когда мы говорим о множестве прямых, то имеем множество, элементами которого являются другие множества. Факт, что некоторый объект a является элементом множества A записывается в виде $a \in A$ ($A \ni a$), что читается как « a лежит в A » (« A имеет a своим элементом»). Аналогично записи $a \notin A$ и $A \not\ni a$ (« a не лежит в A » и « A не содержит a в качестве элемента») означают, что объект a не есть элемент множества A .

Множества называются *равными*, если они состоят из одних и тех же

элементов. Факт, что множества A и B равны записывается так: $A = B$. Если каждый элемент множества A является элементом множества B (запись $A \subset B$ или $B \supset A$), то говорят, что множество A является подмножеством множества B . Из вышесказанного ясно, что $A = B$ в том и только том случае, когда одновременно выполнены оба включения $A \subset B$ и $B \subset A$.

В некоторых книгах для обозначения упомянутых выше включений употребляются символы \subseteq и \supseteq , а символы \subset и \supset означают строгое включение множеств, исключающее их совпадение. Нам нет необходимости разделять случаи строгого и нестрогого включений, поэтому символы \subseteq и \supseteq нами использоваться не будут.

Замечание. Несмотря на внешнюю похожесть записей $B \in A$ и $B \subset A$, их смысл совершенно различный. Так, если A — множество цифр, применяемых в десятичной системе счисления, а B — множество чётных таких цифр, то запись $B \subset A$ будет верной, а запись $B \in A$ — неверной. Легко привести и обратный пример.

Если множество задано перечислением своих элементов, то эти элементы записываются в фигурных скобках через запятую. Так, множества A и B из предыдущего замечания записываются как $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ и $\{0, 2, 4, 6, 8\}$ соответственно. Разумеется, задать таким образом можно очень небольшой процент множеств. Чтобы обойти эту трудность, часто поступают так: вместо описываемого множества A указывают некоторое другое, известное множество, содержащее A , и задают некоторое правило P , которому должны соответствовать те и только те элементы указанного множества, которые и образуют A . Математической записью здесь такова: в фигурных скобках записывается, что переменная принадлежит указанному большему множеству, а далее через двоеточие (или иногда через вертикальную черту) указывают правило P . Например, запись $\{x \in \mathbf{R} : x = \frac{m}{n}, m \in \mathbf{Z}, n \in \mathbf{N}\}$ означает подмножество вещественных чисел, которые можно представить как отношение целого числа к натуральному — из курса средней школы вам известно, что такие числа называются *рациональными*. Иногда, если из контекста ясно, о чём идёт речь, объем-

лющее множество может и не указываться; например, вполне корректна запись $\mathbf{Q} = \{x : x = \frac{m}{n}, m \in \mathbf{Z}, n \in \mathbf{N}\}$.

Упражнение 1. Совпадают ли множества A и B , если

- а) $A = \{1, 2, 3\}, B = \{2, 1, 3\}$;
- б) $A = \{1, 2, 3\}, B = \{1, 2, 3, 1\}$;
- в) $A = \{1, 2, 3\}, B = \{\{1, 2\}, 3\}$?

Существует ряд множеств, для которых исторически закреплены стандартные обозначения. Вот некоторые из них: \mathbf{N} — множество натуральных чисел, \mathbf{Z} — множество целых чисел, \mathbf{Q} — множество рациональных чисел, \mathbf{R} — множество действительных чисел, \mathbf{C} — множество комплексных чисел, \mathbf{R}^2 — множество точек плоскости, \mathbf{R}^3 — множество точек пространства и т. д.

Упражнение 2. Совпадают ли множества A и B , если

- а) $A = \{x \in \mathbf{Z} : x = 3n, n \in \mathbf{Z}\}, B = \{x \in \mathbf{Z} : x = 3n + 6, n \in \mathbf{Z}\}$;
- б) $A = \{x \in \mathbf{Z} : x = 3n, n \in \mathbf{N}\}, B = \{x \in \mathbf{Z} : x = 3n + 6, n \in \mathbf{N}\}$;
- в) $A = \mathbf{Q}, B = \{x \in \mathbf{R} : x = \frac{m}{n}, m \in \mathbf{N}, n \in \mathbf{Z}\}$?

Особо надо сказать о *пустом множестве* (обозначается \emptyset). Так называют множество, которое не содержит ни одного элемента. Пустое множество единственно. Оно является подмножеством любого множества, т. е. для любого A справедливо включение $A \supset \emptyset$. Вопрос о том, является ли то или иное множество пустым или нет, является одним из основных вопросов математики; так вопрос о совместности системы уравнений есть вопрос о непустоте множества её решений. Обратим внимание на то, что множества \emptyset и $\{\emptyset\}$ различны: первое из них не содержит элементов, а второе содержит единственный элемент — пустое множество.

Упражнение 3. Выпишите все подмножества следующих множеств:

- а) $\{t, s\}$, б) $\{2, 5, \text{ъ}\}$, в) $\{\emptyset\}$, г) $\{\emptyset, \{\emptyset\}\}$.

Существует специальное обозначение для множества всех подмножеств некоторого множества A . Оно таково: 2^A . Математическая запись этого множества такая: $2^A = \{B : B \subset A\}$. Например, если множество A — множество цифр, используемых в двоичной системе счисления, (т. е. если $A = \{0, 1\}$), то $2^A = \{\emptyset, \{1\}, \{0\}, A\}$. Обращаем особое внимание, что здесь не идёт речь о возведении числа 2 в степень A , запись 2^A — просто обозначение некоего множества. Объяснение же такому обозначению таково: оказывается (позже мы это докажем), что если множество A состоит из n элементов (n — натуральное число), то множество 2^A состоит из 2^n элементов ровно.

2. Операции над множествами

Основными операциями над множествами являются объединение, пересечение и разность.

Определение 1. Пересечением множеств A и B называют множество, элементами которого являются все элементы, которые принадлежат одновременно множествам A и B и только они.

Определение 2. Объединением множеств A и B называют множество, элементами которого являются все элементы, которые принадлежат хотя бы одному из множеств A и B и только они.

Пересечение и объединение множеств обозначаются символами \cap и \cup соответственно. С учётом этого, определения могут быть переписаны следующим образом:

$$A \cup B = \{x : x \in A \text{ или } x \in B\},$$

$$A \cap B = \{x : x \in A \text{ и } x \in B\}.$$

Для наглядности представления этих операций используются так называемые *диаграммы Венна*, в которых множества, над которыми производится операция, изображаются в виде кругов на плоскости. При этом

элементами множества считаются все точки, лежащие внутри или на границе круга, соответствующего данному множеству, а результат операции закрашивается. На рис. 1 и рис. 2 соответственно представлены диаграммы Венна для объединения и пересечения двух множеств.

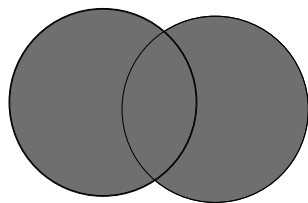


Рис. 1. Объединение
двух множеств

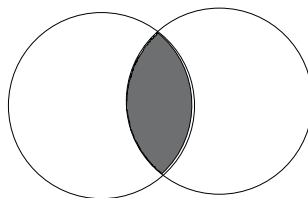


Рис. 2. Пересечение
двух множеств

Операции объединения и пересечения двух множеств легко распространяются на случай произвольного (в том числе и бесконечного) числа множеств:

Определение 1'. *Объединением* семейства множеств A_α называют множество, состоящее из тех и только тех элементов, которые принадлежат хотя бы одному из множеств A_α .

Определение 2'. *Пересечением* семейства множеств A_α называют множество, состоящее из тех и только тех элементов, которые принадлежат одновременно каждому множеству A_α .

Для теоретико-множественной записи этих определений используются так называемые *кванторы*. Их два: *квантор всеобщности* и *квантор существования*.

Квантор всеобщности имеет вид \forall . Это замена русских слов «любой», «каждый», «всякий» и их синонимов. Например, утверждение, что 1 — самое маленькое натуральное число, с помощью квантора всеобщности записывается так:

$$\forall x \in \mathbf{N} \ 1 \leq x.$$

Квантор существования имеет вид \exists . Это замена русских слов «имеется», «существует», «найдётся» и их синонимов. Например, утверждение,

что многочлен $P(x)$ имеет корни, с помощью квантора существования записывается так:

$$\exists a P(a) = 0.$$

В современной математике кванторы общности и существования встречаются постоянно, практически во всех разделах. Неоднократно они будут встречаться и нам. Покажем, например, как с их помощью записываются определения пересечения и объединения совокупности множеств:

$$\bigcup_{\alpha} A_{\alpha} = \{x : \exists \alpha (x \in A_{\alpha})\}, \quad \bigcap_{\alpha} A_{\alpha} = \{x : \forall \alpha (x \in A_{\alpha})\}.$$

Упражнение 4. Что представляют собой множества

$$\text{а) } \bigcup_{n=1}^{\infty} (1/n; n); \quad \text{б) } \bigcap_{n=1}^{\infty} [-n; 1/n]?$$

Отметим, что с пересечением и объединением множеств вы сталкивались в курсе алгебры средней школы, а именно при решении систем и совокупностей. Рассмотрим, например, систему нескольких уравнений от одного переменного x . Для каждого входящего в систему уравнения имеются свои числа x , ему удовлетворяющие; эти числа образуют множество решений уравнения. Системе удовлетворяют числа x , удовлетворяющие каждому из уравнений, т. е. множеством решений системы является пересечение множеств решений входящих в систему уравнений. Аналогично, множество решений совокупности есть объединение множеств решений входящих в совокупность уравнений и неравенств.

Определение 3. Разностью множеств A и B (обозначается $A \setminus B$) называют множество, состоящее из всех элементов, которые принадлежат множеству A , но не принадлежат множеству B (рис. 3). Математическая запись этого определения такова: $A \setminus B = \{x : x \in A, x \notin B\}$.

Упражнение 5. Верно ли, что из того, что $A \setminus B = \emptyset$ следует, что $A = B$?

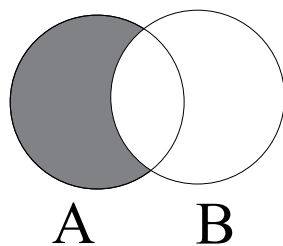


Рис. 3. Разность $A \setminus B$

В отличие от операций объединения и пересечения, операция разности множеств несимметрична: $A \setminus B \neq B \setminus A$. Кроме того, она не распространяется на случай более, чем двух множеств.

Упражнение 6. Для каких пар множеств A , B всё же имеет место равенство $A \setminus B = B \setminus A$?

Часто имеет смысл рассматривать некоторое «максимальное» множество, которое содержит в себе все рассматриваемые. В качестве такого объекта подошло бы, например, множество всех множеств, т. е. множество, элементами которого являются всевозможные множества. Однако оказывается, что такого множества в природе не существует, точнее, *совокупность всех множеств множеством не является* (этот факт в настоящем курсе мы доказывать не будем). Такую совокупность в теории множеств называют *универсум*. Поскольку универсум не есть множество, математика работать с ним не может. Вместо универсума математики рассматривают некий его аналог — *универсальное множество*, т. е. множество, которое содержит все множества, возникающие при решении данной задачи. Как видно, это множество зависит от конкретного раздела математики, в котором решается та или иная задача. Например, в теории делимости универсальным множеством может считаться \mathbf{N} (или \mathbf{Z}), в планиметрии — \mathbf{R}^2 и т. д. Для универсального множества зарезервировано обозначение U . На диаграммах Венна его изображают в виде прямоугольника, внутри которого изображаются кругами все другие множества.

Определение 4. Дополнением множества A (обозначение \overline{A}) называется разность $U \setminus A$.

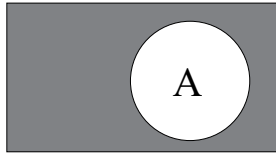


Рис. 4. Дополнение до множества A — это закрашенная область

Ниже приведена диаграмма Венна для операции дополнения (рис. 4).

В некоторых разделах математики (например, в теории меры) полезна операция симметрической разности двух множеств.

Определение 5. Симметрической разностью множеств A и B (обозначение $A \Delta B$) называется множество, состоящее из тех и только тех элементов, которые принадлежат ровно одному из множеств A и B .

На языке теории множеств определение выглядит следующим образом:
 $A \Delta B = (A \setminus B) \cup (B \setminus A)$.

Упражнение 7. Постройте диаграмму Венна для симметрической разности и докажите, что $A \Delta B = (A \cup B) \setminus (A \cap B)$.

3. Алгебра множеств

Рассмотренные в предыдущем пункте операции обладают многочисленными свойствами, которые называются *формулами теории множеств*. Каждая из этих формул выглядит как равенство двух множеств, причём это равенство верно независимо от того, какие конкретные множества обозначены буквами, входящими в формулу. Набор всех таких формул называется *алгеброй множеств*. Ясно, что алгебра множеств состоит из бесконечного числа формул. Некоторые из них считаются основными, поскольку выражают ключевые свойства введённых операций. Все другие формулы могут быть выведены последовательным применением каких-либо из основных формул. Приведём основные формулы теории множеств.

- | | | | |
|---------------------|---------------------|------------------------------------|-------------------------------------|
| 1) $A \cup A = A$; | 2) $A \cap A = A$; | 3) $A \cup \emptyset = A$; | 4) $A \cap \emptyset = \emptyset$; |
| 5) $A \cap U = A$; | 6) $A \cup U = U$; | 5) $\overline{\overline{A}} = A$; | 6) $\overline{\emptyset} = U$. |

Эти свойства совершенно очевидны: они следуют из определения операций объединения, пересечения и дополнения.

$$7) \overline{\overline{A}} = A.$$

Формула 7 имеет специальное название: закон двойного дополнения. Она тоже практически очевидна: \overline{A} — это все элементы, которые не входят в множество A , поэтому $\overline{\overline{A}}$ — это те и только те элементы, которые не входят в множество \overline{A} , а эти элементы и составляют множество A .

$$8) A \cup B = B \cup A; \quad 9) A \cap B = B \cap A;$$

$$10) (A \cap B) \cap C = A \cap (B \cap C); \quad 11) (A \cup B) \cup C = A \cup (B \cup C).$$

Формулы 8 и 9 выражают перестановочность объединяемых (пересекаемых) множеств. Это очевидное свойство операций объединения и пересечения называется *коммутативностью*. Формулы 10 и 11 говорят, что порядок объединения (пересечения) не важен; так называемая, *ассоциативность* этих операций.

Заметим, что понятия коммутативности и ассоциативности вам известны из курса средней школы: это переместительный и сочетательный закон, которому удовлетворяют также обычные операции сложения и умножения чисел. Если в формулах 8 — 11, а также в формулах 3 и 4 формально заменить множества на числа, символ \cup на $+$, символ \cap на \cdot , а пустое множество считать нулём, мы получим хорошо известные со школы правила арифметики. Эта закономерность привела к тому, что операцию объединения двух множеств часто называют теоретико-множественным сложением, а пересечения — теоретико-множественным умножением. Однако далеко не все формулы теории множеств имеют такой арифметический аналог (например, формула 1 при приведённой выше трактовке получила бы вид $a + a = a$ — равенство совершенно неверное).

$$12) (A \cap B) \cup C = (A \cup C) \cap (B \cup C); \quad 13) (A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

Эти две формулы выражают два свойства *дистрибутивности*. Они не столь очевидны, как предыдущие, и нуждаются в доказательстве, или хотя бы пояснении. Особенно это касается формулы 12, так как другая формула имеет очевидную числовую аналогию: равенство $(a + b)c = ac + bc$. А вот свойства $ab + c = (a + c)(b + c)$ в алгебре нет. Проиллюстрировать эти формулы (равно как и все предыдущие) можно диаграммами Венна. Вот как выглядит такая иллюстрация для формулы 12 (рис. 5).

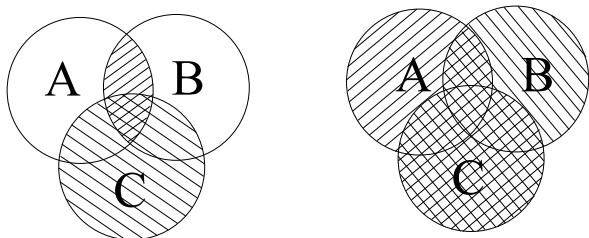


Рис. 5. Иллюстрация формулы 12

На левой диаграмме изображена левая часть формулы. В одном направлении заштриховано множество $A \cap B$, в перпендикулярном — множество C . Левая часть — это всё, что заштриховано хотя бы в одну сторону. Аналогично, на правой диаграмме изображено правое

множество формулы 12. Различными штриховками отмечены пересекаемые множества, поэтому в правой части формулы находится множество, покрытое обеими штриховками сразу. Нетрудно видеть, что конечные множества на обеих диаграммах одинаковы.

Строгое же доказательство (заметим, что так можно доказать любую формулу теории множеств) заключается в том, что доказываются два включения \supset и \subset , то есть, сперва берётся произвольный элемент из множества в левой части формулы и показывается, что он лежит в правой, а потом наоборот. Вот как выглядит доказательство формулы 13:

\supset . $\forall x \in (A \cap C) \cup (B \cap C) \Rightarrow x \in (A \cap C)$ или $x \in (B \cap C)$. Пусть, без ограничения общности, $x \in (A \cap C)$. Тогда $x \in A$ и $x \in C$. Из первого включения следует (так как $A \cup B \supset A$), что $x \in A \cup B$. Учитывая, что $x \in C$ получаем $x \in (A \cup B) \cap C$.

\subset . $\forall x \in (A \cup B) \cap C \Rightarrow x \in (A \cup B)$ и $x \in C \Rightarrow x \in A$ или $x \in B$ и $x \in C$. Пусть $x \in A$ (случай $x \in B$ аналогичен). Тогда (так как $x \in C$)

имеем $x \in A \cap C$ и $x \in (A \cap C) \cup (B \cap C)$.

Следующие две формулы называют *законами поглощения*:

$$14) (A \cap B) \cup A = A; \quad 15) (A \cup B) \cap A = A.$$

Упражнение 8. Проиллюстрируйте законы поглощения диаграммами Венна.

Завершают список основных формул теории множеств *законы де Моргана*:

$$16) \overline{(A \cap B)} = \bar{A} \cup \bar{B}; \quad 17) \overline{(A \cup B)} = \bar{A} \cap \bar{B};$$

$$18) A \setminus \bigcup_{\alpha} B_{\alpha} = \bigcap_{\alpha} (A \setminus B_{\alpha}); \quad 19) A \setminus \bigcap_{\alpha} B_{\alpha} = \bigcup_{\alpha} (A \setminus B_{\alpha}).$$

Заметим, что если законы 16 и 17 могут быть изображены на диаграммах, то законы 18 и 19 не имеют графической картинки, так как в них присутствует бесконечное число множеств. Их доказательство может быть проведено только аналитическим путём.

Упражнение 9. Проведите доказательство формул 16–19.

Упражнение 10. Докажите, что симметричная разность — есть операция ассоциативная, т. е. $A \Delta (B \Delta C) = (A \Delta B) \Delta C$.

4. Декартово произведение множеств

Определение 6. Упорядоченной парой (a, b) называется множество $\{a, \{a, b\}\}$, упорядоченной тройкой (a, b, c) — множество $(a, (b, c))$, вообще для произвольного натурального n упорядоченной n -й (a_1, a_2, \dots, a_n) называется множество $(a_1, (a_2, \dots, a_{n-1}))$ (определение дано по индукции).

Несмотря на некоторую абстрактность и неудобочитаемость последнего определения, оно имеет весьма простой смысл: упорядоченная пара — это двухэлементное множество, в котором указано, какой элемент считать первым, а какой — вторым. Аналогичный смысл имеют упорядоченные тройки, четвёрки и т. д.

Определение 7. *Декартовым произведением* множеств A и B (обозначается $A \times B$) называют множество всех упорядоченных пар, в которых первый элемент принадлежит множеству A , а второй — множеству B .

На формальном языке определение выглядит так:

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

Аналогично определяется декартово произведение любого конечного семейства множеств:

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) : \forall i = 1, 2, \dots, n \ a_i \in A_i\}.$$

С декартовым произведением вы уже встречались в школе, когда проходили метод координат (не важно, на плоскости или в пространстве). Собственно говоря, координатную плоскость надлежит трактовать как декартово произведение двух прямых (абсциссы и ординаты), а пространство — как декартово произведение трёх прямых (добавляется аппликата). Именно эта трактовка и позволяет обозначать плоскость и пространство символами \mathbf{R}^2 и \mathbf{R}^3 . Другой пример декартова произведения представляет собой шахматная доска: известно, что поля шахматной доски обозначаются парой буква — цифра (латинскими буквами с a по h обозначены вертикали доски, цифрами от 1 до 8 — её горизонталы; символом $d2$ обозначено поле, находящееся на пересечении вертикали « d » с горизонталью «2» и т. д.).

В общем случае $A \times B \neq B \times A$ (приведите самостоятельно пример таких множеств A и B), т. е. декартово произведение не коммутативно, однако вполне очевидна его ассоциативность: $A \times (B \times C) = (A \times B) \times C$, что позволяет обойтись без скобок при написании таких произведений.

Если же все перемножаемые множества равны, то представляется разумным использовать степенную запись (A^2 вместо $A \times A$ и т. п.), как это всегда и делается.

Упражнение 11. Справедливы ли формулы:

а) $(A \cup B) \times C = (A \times C) \cup (B \times C)$;

б) $(A \cap B) \times C = (A \times C) \cap (B \times C)$;

в) $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$?

Сформулируйте и проверьте аналогичные формулы с теоретико-множественными операциями во втором сомножителе.

ЛЕКЦИЯ II. ФУНКЦИИ И ГРАФИКИ

1. Основные определения

Определение 8. Пусть даны два непустых множества A и B . *Отображением f или функцией*, действующей из множества A в множество B , называют правило, которое каждому элементу множества A ставит в соответствие единственный элемент множества B . Записывается это так: $f : A \rightarrow B$.

Если при этом множество B — это некоторое подмножество числовой прямой, то говорят, что функция f *вещественная* или *вещественнозначная*.

Факт, что функция f ставит в соответствие элементу $a \in A$ элемент $b \in B$ записывается известным со школы образом, а именно: $f(a) = b$.

Функции, изучаемые вами в средней школе, как правило были *числовыми* функциями, т. е. множество A всегда предполагалось подмножеством числовой прямой. Реально часто приходится иметь дело с функциями не числовыми, например, когда мы говорим о росте человека, мы имеем

ввиду функцию, действующую из множества всех людей в числовую прямую. Формула пройденного пути при равномерном движении $S = vt$, где v — скорость движущегося тела, а t — время движения, может быть рассмотрена как функция двух переменных $S = S(v, t)$, т. е. $S : \mathbf{R}^2 \rightarrow \mathbf{R}$. Нетрудно также привести пример функций, не являющихся вещественнозначными, но мы этого делать не будем. Хотя вся излагаемая ниже теория справедлива для всех типов функций, мы будем приводить в качестве примеров только числовые функции.

Грубо говоря, функция — это правило соответствия между двумя множествами. Однако следует иметь в виду, что функция — это не только само правило, это ещё и два множества A и B . Изменив одно из них, мы изменим и функцию. Например, если $f(x) = x^2$, то попарно различны отображения $f : \mathbf{R} \rightarrow \mathbf{R}$, $f : \mathbf{R} \rightarrow [0; +\infty)$ и $f : [0; +\infty) \rightarrow \mathbf{R}$. В самом деле, второе из этих отображений обладает свойством (далее мы назовём такие функции сюръективными), что в каждую точку $b \in B$ какая-нибудь точка, да отображается, в то время как первая и третья функции этому условию не удовлетворяют. Первая функция от третьей отличается, например, тем, что третья функция монотонно возрастает (если $x_1 > x_2$, то $f(x_1) > f(x_2)$), про первую функцию этого сказать нельзя.

Таким образом, функция $f : A \rightarrow B$ — это упорядоченная тройка (A, B, f) , и при различных A или B получаются разные функции.

Обратим внимание, что в отличие от изучаемых в школе функций, мы всегда полагаем, что всякая функция $f : A \rightarrow B$ определена на всём множестве A , поэтому вопрос об области определения не стоит. Имеющееся противоречие со школьной программой (все вы знаете, что в алгебре бывают задачи, в которых стоит задание типа «найдите область определения данной функции») кажущееся. Дело в том, что в школьной алгебре (а часто и в математическом анализе) множество A не указывается, просто предполагается, что $A \subset \mathbf{R}$ и делается следующее допущение: функция определена на максимально допустимом подмножестве числовой прямой. Например, запись $f(x) = \sqrt{x}$ подразумевает функцию $f : [0; \infty) \rightarrow \mathbf{R}$. Тогда вопрос об области определения становится действительно актуальным.

Также (если рассматриваются вещественные функции) часто не указывается множество B . Если не оговорено противное, полагается, что B — вся числовая прямая. В ряде случаев это действительно удобное соглашение, ведь в определении функции вовсе не требуется, чтобы отображение f принимало все значения из B .

Что является действительно важным в понятии функции $f : A \rightarrow B$, так это то, что каждому элементу $a \in A$ отображение ставит в соответствие *единственное* значение $f(a)$. Если этого нет, то мы имеем дело не с функцией, а с более сложным математическим объектом: *многозначным отображением*. В настоящем цикле лекций мы не будем касаться многозначных отображений.

Важнейшими понятиями, связанными с функциями, являются *образ* и *прообраз*.

Определение 9. *Образом множества $X \subset A$ при отображении $f : A \rightarrow B$ называется множество $\{b \in B : b = f(a) \text{ для некоторого элемента } a \in X\}$. Это множество обозначается $f(X)$.*

Определение 10. *Прообразом (полным) множества $Y \subset B$ при отображении $f : A \rightarrow B$ называется множество $\{a \in A : f(a) \in Y\}$. Это множество обозначается $f^{-1}(Y)$.*

Если из контекста ясно, о каком отображении идёт речь, то в определении образа и прообраза ссылки на отображение не делают, говоря, например: «образ отрезка», «прообраз точки» и т. п.

Иногда имеет смысл говорить о частичном прообразе точки (или реже множества), т. е. о некотором подмножестве прообраза этой точки (множества). Здесь иногда (когда смысл не допускает двоякого толкования) также слово частичный может быть опущено. Например, вполне можно сказать, что точка 0 является прообразом точки 1 при отображении $f(x) = \cos x$. Понятно, что здесь идёт речь о том факте, что $\cos 0 = 1$, и не больше. (Вам уже известно, что полным прообразом 1 в этом случае является множество

$\{2\pi n : n \in \mathbf{Z}\}.$)

Обратите внимание на запись полного прообраза: $f^{-1}(Y)$. Её ни в коем случае нельзя толковать, как $\frac{1}{f(Y)}$. Во-первых, в случае не вещественной функции последняя запись просто бессмысленна: ведь $f(Y)$ не число, и разделить на него единицу нельзя. Во-вторых, даже если функция числовая, и даже если полный прообраз точка, значения $f^{-1}(Y)$ и $\frac{1}{f(Y)}$ чаще всего различны. Запись $f^{-1}(Y)$ для полного прообраза связана с понятием обратного отображения, о котором будем говорить чуть позже. В данный момент просто надлежит понять, что $f(X)$ и $f^{-1}(Y)$ — это просто некоторые множества ($f(X) \subset B$, $f^{-1}(Y) \subset A$), и над ними можно проделывать те самые операции, о которых шла речь в первой лекции: пересекать, объединять, брать дополнение и т. д.

Упражнение 12. Докажите следующие теоретико-множественные соотношения.

$$\begin{array}{ll} \text{а) } f(\bigcup_{\alpha} A_{\alpha}) = \bigcup_{\alpha} f(A_{\alpha}); & \text{б) } f(\bigcap_{\alpha} A_{\alpha}) \subset \bigcap_{\alpha} f(A_{\alpha}); \\ \text{в) } f^{-1}(\bigcup_{\alpha} B_{\alpha}) = \bigcup_{\alpha} f^{-1}(B_{\alpha}); & \text{г) } f^{-1}(\bigcap_{\alpha} B_{\alpha}) = \bigcap_{\alpha} f^{-1}(B_{\alpha}). \end{array}$$

С понятием функции тесным образом связано понятие её графика.

Определение 11. Графиком функции $f : A \rightarrow B$ (обозначение $\Gamma(f)$) называют множество $\{(a, b) \in A \times B : b = f(a)\}$.

Заметим, что понятие графика существует для любой функции (не обязательно числовой), только график при этом выглядит не как линия на координатной плоскости, а более сложным образом. Так, если рассматривать функцию $f : \mathbf{R}^2 \rightarrow \mathbf{R}$, то её график — это подмножество трёхмерного пространства \mathbf{R}^3 , обычно поверхность. В более сложных случаях график функции ещё менее нагляден.

Также заметим, что даже в случае числовых функций график далеко не всегда представляет из себя линию. Например, рассмотрим функцию

Дирихле: $f(x) = \begin{cases} 1, & \text{если } x \in \mathbf{Q} \\ 0, & \text{если } x \notin \mathbf{Q} \end{cases}$. Если попытаться изобразить её график, то на плоскости получатся две прямые $y = 0$ и $y = 1$ — настолько близко друг к другу расположены рациональные и иррациональные числа. Однако это — оптический обман, мираж. Реально каждая из прямых — частей графика — не сплошная, в ней много «дыр». Насколько много, что каждая вертикальная прямая проходит через дыру либо на одной прямой, либо на другой. Ясно, что график такой функции не совсем нагляден. Однако он есть.

Разумеется, не всякое подмножество множества $A \times B$ может выступать графиком какого-либо отображения. Из определения функции ясно, что графиком функции может быть только такое подмножество M декартова произведения $A \times B$, что для любого элемента $a \in A$ существует ровно один элемент $b \in B$, удовлетворяющий условию $(a, b) \in M$. Наоборот, всякое подмножество $M \subset A \times B$, удовлетворяющее этому свойству, определяет функцию, график которой есть множество M . Графически (в случае функции $f : \mathbf{R} \rightarrow \mathbf{R}$) это означает, что всякая вертикальная прямая может пересекать график не более чем по одной точке. Таким образом, например, окружность не является графиком никакой функции, прямая является графиком функции тогда и только тогда, когда она не вертикальна и т. д.

Упражнение 13. Как по графику функции $f : \mathbf{R} \rightarrow \mathbf{R}$ определить:

- а) образ множества $M \subset R$;
- б) прообраз множества $M \subset R$;
- в) область определения функции f ;
- г) область значения функции f ?

Связь между функцией и её графиком настолько велика, что иногда эти объекты отождествляют. Тогда всякая функция $f : A \rightarrow B$ есть просто подмножество декартова произведения $A \times B$, т. е. вместо изучения функций можно изучать множества. Иногда это действительно удобно, и

приводит к нетривиальным результатам, однако сущность функции (как соответствия множеств) и наглядность при таком подходе теряются, поэтому мы так поступать не будем.

2. Действия над функциями. Простейшие числовые функции и их графики

В этом подразделе лекции мы будем рассматривать исключительно вещественнозначные функции, поскольку если они таковыми не являются, то невозможно производить над ними обычные арифметические действия. Более точно, для того, чтобы функции можно было, например, сложить необходимо и достаточно, чтобы операция сложения была определена на множестве значений этих функций. Если, например, две функции действуют из одного и того же множества в пространство векторов, то их можно складывать (поскольку можно складывать векторы), а делить нельзя (операции деления на векторах нет). Кроме того, мы будем полагать, что все функции действуют из одного и того же множества (чаще всего, это будет некоторое подмножество числовой прямой).

Определение 12. Пусть $f : A \rightarrow \mathbb{R}$ и $g : A \rightarrow \mathbb{R}$ — две произвольные вещественнозначные функции. Суммой $f + g$ (соответственно разностью $f - g$, произведением fg , частным $\frac{f}{g}$) называют функцию $(f + g) : A \rightarrow \mathbb{R}$, действующую по правилу: $(f + g)(x) = f(x) + g(x)$ $\left((f - g)(x) = f(x) - g(x), (fg)(x) = f(x) \cdot g(x), \frac{f}{g}(x) = \frac{f(x)}{g(x)} \right)$ для всех $x \in A$. (Функция частного $\frac{f}{g}$ определена для тех и только тех $x \in A$, в которых $g(x) \neq 0$.)

Замечание 1. Иногда приходится рассматривать арифметические операции над функциями, область определения которых различна. В этом случае результирующая функция определяется на *пересечении* областей определения. При этом можно считать (мы так и будем делать), что функции,

над которыми проводится алгебраическое действие, изначально определены только на этом пересечении. Так, при сложении функций $f(x) = \sqrt{x}$ и $g(x) = \frac{2}{x^2 - 1}$ (определённых на множествах $[0; +\infty)$ и $\mathbf{R} \setminus \{-1, 1\}$ соответственно) получается функция $\sqrt{x} + \frac{2}{x^2 - 1}$, определённая на множестве $[0; 1) \cup (1; +\infty)$; мы договорились считать, что операции сложения подверглись сужения функций f и g на множество $[0; 1) \cup (1; +\infty)$.

Замечание 2. Поскольку любое действительное число t можно считать постоянной функцией (принимаящей во всех точках значение, равное числу t), то для любой вещественнозначной функции f и любого действительного числа t определены также функции $f + t$, $f - t$ и tf , а если $t \neq 0$, то и функция $\frac{f}{t}$. Все эти функции определены на области определения функции f .

Упражнение 14. Пусть $f(x) = 2^x$ и $g(x) = \frac{1}{x}$. Найдите, чему равны

а) $(f \cdot g)(4)$; б) $(f - g)(-1)$; в) $(2f)(0)$; г) $(3f + 5g)(2)$.

Упражнение 15. Пусть $f(x) = \operatorname{tg} x$ и $g(x) = \operatorname{ctg} x$. Равны ли функции $h_1 = f \cdot g$ и функция $h_2 \equiv 1$? Постройте графики функций h_1 и h_2 .

Теперь обратимся к стандартным функциям, изучающимся в программе средней школы. Разумеется, все эти функции действуют из некоторого подмножества числовой прямой в числовую прямую. Вот полный список этих функций:

- 1) линейная функция $y = kx + b$, где $k \neq 0$;
- 2) квадратичная функция $y = ax^2 + bx + c$, где $a \neq 0$;
- 3) функция обратной пропорциональности $y = \frac{a}{x}$, где $a \neq 0$;
- 4) степенная функция $y = x^n$, где $n \in \mathbf{N}$;
- 5) функция корня n -ой степени $y = \sqrt[n]{x}$, где $n \in \mathbf{N}$;
- 6) показательная функция $y = a^x$, где $a > 0$, $a \neq 1$;

- 7) логарифмическая функция $y = \log_a x$, где $a > 0$, $a \neq 1$;
- 8) основные тригонометрические функции $y = \sin x$, $y = \cos x$, $y = \operatorname{tg} x$ и $y = \operatorname{ctg} x$.

Эти функции (а также функции, обратные основным тригонометрическим на соответствующем промежутке, о которых мы подробно будем говорить позже) называются *простейшими*. Каждый из вас обязан хорошо представлять поведение этих функций и их графики, а также знать их свойства. В целях экономии времени и места мы не приводим здесь этих фактов; желающие вспомнить их могут обратиться к учебникам алгебры за курс средней школы.

Вам, конечно, знакомо уравнение окружности с центром в точке $(a; b)$ радиуса R : $(x - a)^2 + (y - b)^2 = R^2$. Это уравнение не задаёт однозначно функции, однако, выразив из него переменную y , мы можем получить две функции: $y = b + \sqrt{R^2 - (x - a)^2}$ и $y = b - \sqrt{R^2 - (x - a)^2}$. Первое из них задаёт верхнюю полуокружность, а второе — нижнюю.

3. Преобразования графиков функций

В случае функций, действующих из $X \subset \mathbf{R}$ в \mathbf{R} зачастую можно проследить, как меняется график функции при применении к ней арифметической операции. Точнее, имеет место теорема.

Теорема 1. Пусть $X \subset \mathbf{R}$, $f : X \rightarrow \mathbf{R}$ и $a > 0$. Тогда:

- а) график функции $f(x) + a$ получается из графика функции f сдвигом на a единиц вверх (вдоль оси ординат);
- б) график функции $f(x) - a$ получается из графика функции f сдвигом на a единиц вниз (вдоль оси ординат);
- в) график функции $f(x + a)$ получается из графика функции f сдвигом на a единиц влево (вдоль оси абсцисс);
- г) график функции $f(x - a)$ получается из графика функции f сдвигом на a единиц вправо (вдоль оси абсцисс);

д) график функции $f(ax)$ получается из графика функции f сжатием в a раз вдоль оси абсцисс (при $0 < a < 1$ реально получается растяжение);

е) график функции $af(x)$ получается из графика функции f растяжением в a раз вдоль оси ординат (при $0 < a < 1$ реально получается сжатие);

ё) график функции $-f(x)$ получается из графика функции f симметричным отображением относительно оси абсцисс;

ж) график функции $f(-x)$ получается из графика функции f симметричным отображением относительно оси ординат;

з) график функции $|f(x)|$ получается из графика функции f симметричным отображением относительно оси абсцисс той части графика, которая лежит ниже этой оси (часть графика функции f , лежащая выше оси абсцисс, остаётся неизменной);

и) график функции $f(|x|)$ получается из графика функции f следующей операцией: часть графика, лежащая левее оси ординат заменяется на образ остальной части графика при симметрии относительно оси ординат (часть графика функции f , лежащая правее оси ординат не изменяется).

Доказательство. Все пункты теоремы доказываются одинаково и независимо друг от друга. Докажем, например, пункт (а).

Пусть точка $M(x_M, y_M)$ лежит на графике функции f . Это выполняется тогда и только тогда, когда $y_M = f(x_M)$. Последнее условие равносильно тому, что $y_M + a = f(x_M) + a = (f + a)(x_M)$, а это эквивалентно тому, что точка $N(x_M, y_M + a)$ лежит на графике функции $f + a$. Остаётся заметить, что точка N получается из точки M сдвигом вдоль оси Oy вверх на a единиц. Пункт (а) доказан.

Доказательство пунктов (б) — (и) мы оставляем читателю. Как уже говорилось, оно аналогично доказательству пункта а).

Разумеется, операции сдвига (пункты (а) — (г) теоремы) и осевой симметрии (пункты (ё) и (ж)) гораздо нагляднее операций растяжения и сжатия (пункты (д) и (е)); неудивительно, что в стандартном курсе средней

школы сжатие и растяжение графиков даются только для тригонометрических функций: там указанные преобразования имеют чётко выраженный физический смысл: умножение на константу всей функции означает увеличение амплитуды колебаний (если модуль константы меньше 1, то уменьшение), а умножение на константу аргумента — уменьшение частоты (увеличение при $0 < a < 1$).

Преобразования графиков, описываемые пунктами (з) и (и), не вполне очевидны, и, возможно, разумно будет их пояснить.

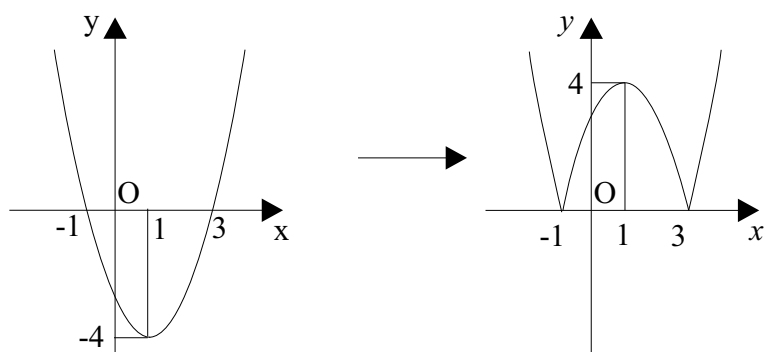


Рис. 1. Взятие модуля от функции

Взятие модуля от всей функции (операция (з)) можно трактовать так: представим себе, что график функции сделан из материала, который очень хорошо прилипает к верхней полуплоскости и очень плохо к нижней. Согнём плоскость вдоль оси абсцисс и совместим полуплоскости друг с другом, а потом развернём обратно. В силу свойства материала, из которого сделан график, он весь приклеится к верхней полуплоскости; то, что получится при этом и является графиком функции $|f|$. Например, построим график функции $y = |x^2 - 2x - 3|$ (рис. 1).

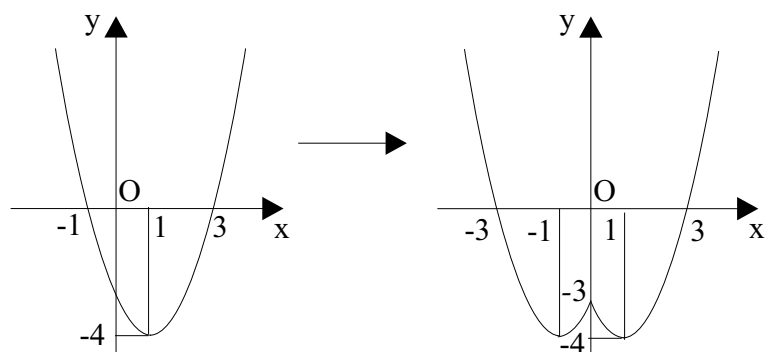


Рис. 2. Взятие модуля от аргумента

Взятие модуля от всей функции (операция (з)) можно трактовать так: представим себе, что график функции сделан из материала, который очень хорошо прилипает к верхней полуплоскости и очень плохо к нижней. Согнём

Взятие модуля от аргумента можно трактовать так: представим себе, что вся ось ординат является вертикально стоящим зеркалом, трюмо, в которое мы смотрим с положительного конца оси Ox .

Тогда мы не видим части графика, которая «за зеркалом»; вместо неё мы видим изображение той

части, которая «до зеркала», эта часть раздвоилась, у неё появился двойник там, «в зазеркалье». То что мы видим (часть графика справа от оси абсцисс и её двойника) и есть график $f(|x|)$. Именно так построен график функции $y = x^2 - 2|x| - 3$ (рис. 2).

Последовательно используя различные пункты теоремы 1 (возможно, некоторые не по одному разу), можно строить графики достаточно сложных функций. Например, легко построить график функции $y = 0,5\sqrt{3 - |x - 2|} + 2$. Поясним, как это делается. Рассмотрим функцию $y_1 = 0,5\sqrt{3 - |x - 2|}$. Согласно пункту (а) теоремы 1, $\Gamma(y)$ получается из графика $\Gamma(y_1)$ сдвигом вверх на 2 единицы. Далее, $\Gamma(y_1)$ получается из графика $\Gamma(y_2)$, где $y_2 = \sqrt{3 - |x - 2|}$, растяжением по вертикали в 0,5 раз, т. е. сжатием в 2 раза (операция (е)). Следующие шаги таковы: рассмотрим функции $y_3 = \sqrt{3 - |x|}$, $y_4 = \sqrt{3 - x}$, $y_5 = \sqrt{x + 3}$, $y_6 = \sqrt{x}$. Каждый график этой последовательности строится из последующего одной из рассмотренных операций, а последний график известен вам со школы. Теперь алгоритм ясен: строим график y_6 , сдвигаем его на 3 единицы влево, отображаем симметрично оси ординат, строим «зеркальный образ» (операция (и)), сдвигаем ещё на 2 единицы вправо, затем сжимаем в 2 раза вдоль оси ординат и, наконец, сдвигаем полученное на 2 единицы вверх (получаемые графики изображены на рис. 3). График функции y построен.

При построениях такого типа следует иметь в виду два момента: 1) не всякий график можно построить такими преобразованиями (например, график функции $y = x + \sqrt{x}$ не построить); 2) «упрощая» вид графика следует следить за тем, чтобы переход к более сложному был возможен, например, в предыдущем примере мы не могли сразу (вместо графика y_1) перейти к графику функции $y = \sqrt{3 - |x - 2|} + 2$, поскольку мы не умеем умножать графически часть функции на константу (правда, если бы нам хотелось, мы могли бы перейти к графику функции $2y = \sqrt{3 - |x - 2|} + 4$, что тоже приводит к решению).

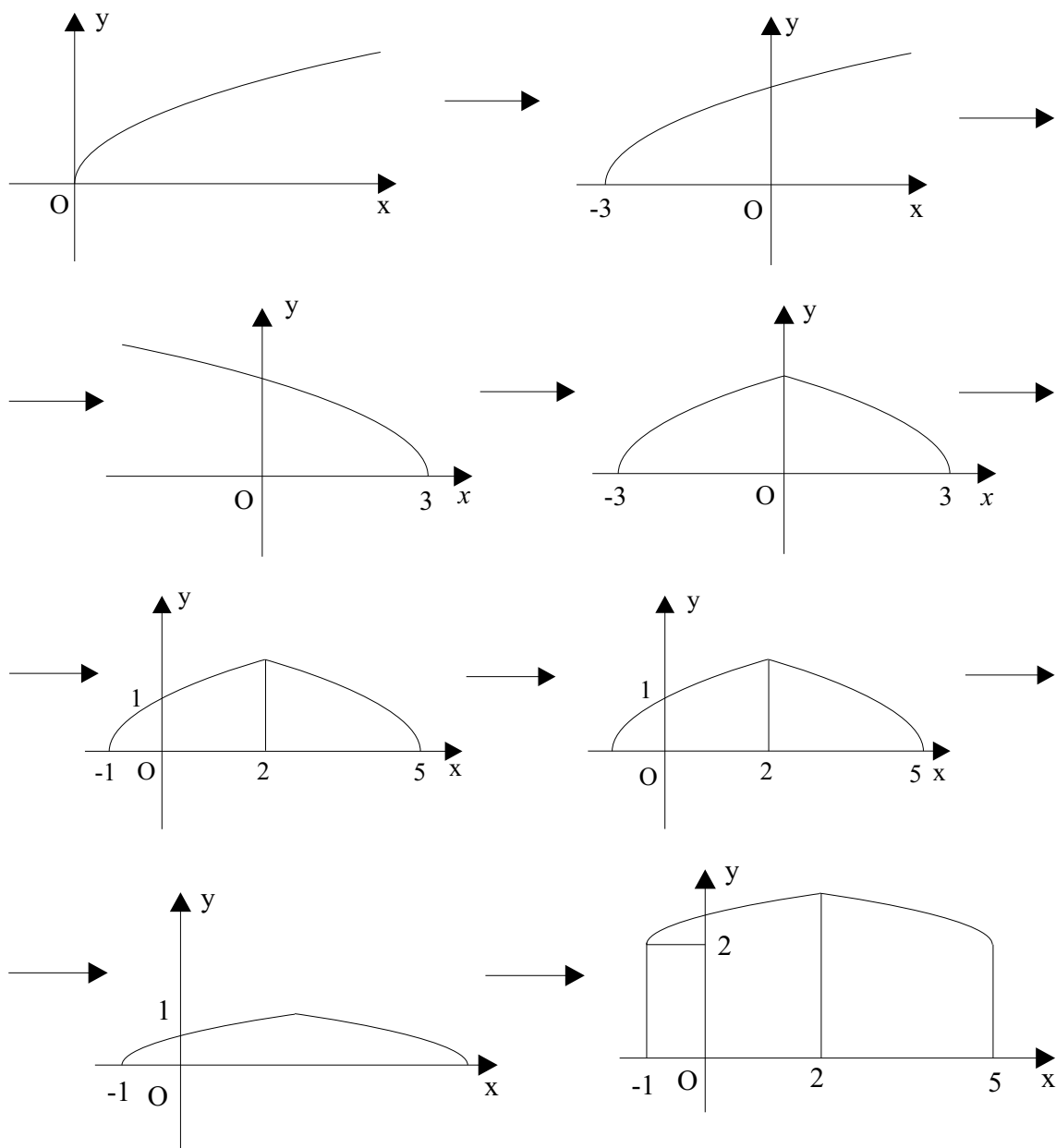


Рис. 3. Пошаговое построение графика функции $y = 0,5\sqrt{3 - |x - 2|} + 2$

Упражнение 16. Используя теорему 1 и графики простейших функций, постройте графики следующих отображений: а) $y = \lg(2 - x)$; б) $y = 3 \cos\left(2x - \frac{\pi}{3}\right)$; в) $y = |2^{|x|} - 5|$; г) $y = \sqrt{4|x| - x^2}$.

В завершении этого пункта покажем, как с помощью преобразований анализируется *дробно-линейная* функция $y = \frac{ax + b}{cx + d}$, где $c \neq 0$, $\frac{a}{c} \neq \frac{b}{d}$ и построим её график.

Прежде всего произведём равносильные преобразования:

$$\frac{ax+b}{cx+d} = \frac{a}{c} \cdot \frac{x+\frac{b}{a}}{x+\frac{d}{c}} = \frac{a}{c} \left(1 + \frac{\frac{b}{a} - \frac{d}{c}}{x+\frac{d}{c}} \right) = \frac{a}{c} + \frac{\frac{bc-ad}{c^2}}{x+\frac{d}{c}}.$$

(Мы при преобразованиях считали, что $a \neq 0$, однако легко убедиться, что заключительная формула также остаётся верной и в случае $a = 0$.)

Теперь видно, что наша функция представляет собой функцию обратной пропорциональности $y = \frac{\frac{bc-ad}{c^2}}{x}$ (обратите внимание, что числитель этой дроби не равен нулю по определению дробно-линейной функции), график которой сдвинут влево на $\frac{d}{c}$ и вверх на $\frac{a}{c}$ (в случае, когда эти числа отрицательны, сдвиг производится соответственно вправо и или вниз). Итак, график дробно-линейной функции — это гипербола. Чтобы найти уравнения её асимптот, применим следующие рассуждения:

вертикальная асимптота $x = x_0$ определяется из условия, что в этой точке функция не определена, т. е. $cx_0 + d = 0$, поэтому её уравнение $x = -\frac{d}{c}$;

горизонтальная асимптота $y = y_0$ обладает тем свойством, что если положить x «очень большим» по модулю числом, то значение функции будут близки к числу y_0 . Заметим, что при больших значениях переменной x дробь $\frac{\frac{bc-ad}{c^2}}{x+\frac{d}{c}}$ является «очень маленьким» по модулю числом, практически

нулём, поэтому функция $y = \frac{a}{c} + \frac{\frac{bc-ad}{c^2}}{x+\frac{d}{c}}$ становится «почти равной» числу $\frac{a}{c}$, т. е. $y_0 = \frac{a}{c}$, а уравнение горизонтальной асимптоты $y = \frac{a}{c}$.

Упражнение 17. Постройте схематически график функции
а) $y = \frac{x}{x+2}$; б) $y = \frac{5x-4}{3x-5}$. Напишите уравнения горизонтальной и вертикальной асимптот.

ЛЕКЦИЯ III. ОБРАТНАЯ ФУНКЦИЯ И ЕЁ ГРАФИК

1. Композиция функций

Определение 13. Пусть даны функции $f : X \rightarrow Y$ и $g : Y \rightarrow Z$. Композицией функций f и g (обозначение $g \circ f$) называется функция, действующая из множества X во множество Z по следующему правилу: $(g \circ f)(x) = g(f(x))$ для всех $x \in X$.

Таким образом, композиция функций по сути есть последовательное применение этих функций. В некоторых учебниках композицию функций называют *суперпозицией*, а в математическом анализе применяют термин *сложная функция*. Вы должны хорошо понимать, что все эти понятия являются синонимами.

В определении 13 имеется неточность, Впрочем, не очень существенная. Дело в том, что функция f может действовать не на всё множество Y , а только на его часть $Y_1 \subset Y$. Поэтому для того, чтобы композиция была определена, достаточно, чтобы функция g была определена не на всём множестве Y , а только на его части — множестве Y_1 . Поэтому более точным было бы рассматривать функцию $g : f(X) \rightarrow Z$. Собственно, так всегда и делается, но для краткости определение обычно дают в том виде, в котором оно приведено.

В математическом анализе, как уже отмечалось выше, часто не указывают область определения функции, подразумевая, что она определена на максимально возможном подмножестве числовой прямой. В этом случае представляет интерес вопрос об области определения композиции двух числовых функций. Совершенно очевидно, что композиция функций не определена в тех точках $x \in X$, в которых не определена функция f , а также в тех точках x , в которых точка $f(x)$ не принадлежит области определения функции g . В остальных же точках композиция определена. Нетрудно понять, что область определения функции $g \circ f$ представляет из себя прообраз области определения функции g при отображении f . Если мы примем обозначение $D(f)$ для области определения функции f ,

этот вывод можно записать формулой $D(g \circ f) = f^{-1}(D(g))$.

Аналогично обстоит дело с множеством значений композиции функций f и g . Легко видеть, что множество значений — это множество $g(f(X))$.

Особый интерес представляет случай, когда множества X , Y и Z являются подмножествами одного и того же множества (например, числовой прямой). В этом случае можно рассматривать композицию функции с собой ($f \circ f$), можно рассматривать одновременно композиции $f \circ g$ и $g \circ f$, можно применять несколько композиций подряд (например, можно рассмотреть функцию $f \circ g \circ g \circ f$) и т. д. Возникает богатое поле деятельности для различных математических исследований.

Упражнение 18. Пусть $f(x) = \cos x$, $g(x) = x^2$ (обе функции действуют из \mathbf{R} в из \mathbf{R}). Что собой представляют функции $f \circ g$ и $g \circ f$? Найдите их области определения и множества их значений.

2. Понятие функции, обратной к данной

Перейдём к понятию обратной функции.

Определение 14. Пусть дана функция $f : X \rightarrow Y$. Тогда функция $g : Y \rightarrow X$ называется *обратной* к функции f , если одновременно выполнены два условия:

- 1) $g(f(x)) = x$ для любой точки $x \in X$;
- 2) $f(g(y)) = y$ для любой точки $y \in Y$.

Например, рассмотрим функцию $f : \mathbf{R} \rightarrow \mathbf{R}$, действующую по правилу $f(x) = 2x - 3$. Для неё обратной функцией будет функция $g(x) = \frac{x+3}{2}$. В самом деле, для всех $x \in \mathbf{R}$ выполнены равенства

$$g(f(x)) = \frac{f(x) + 3}{2} = \frac{2x - 3 + 3}{2} = x \quad \text{и}$$

$$f(g(x)) = 2 \cdot g(x) - 3 = 2 \cdot \frac{x+3}{2} - 3 = x.$$

Заметим, что функция f в свою очередь является обратной к функции g . Это не случайно. Как видно из определения 14, если функция f имеет обратной функцию g , то функция g имеет обратной функцию f . Часто такие функции f и g называют *взаимно обратными*.

Обратную функцию к функции f обозначают символом f^{-1} . Это обозначение никак не согласуется с понятием отрицательной степени, которое вам известно из школы: если a — число, то через a^{-1} обозначается число $\frac{1}{a}$, в то время, как $f^{-1} \neq \frac{1}{f}$ (приведённый только что пример хорошо иллюстрирует этот факт). Поэтому будьте внимательны при работе с этими обозначениями.

Теперь вам должно быть ясно, почему прообраз множества A при отображении f обозначается $f^{-1}(A)$. По сути эта запись означает образ множества A при отображении, обратном к f , отображении f^{-1} . Впрочем, так трактовать прообраз множества можно не всегда; чтобы понять, в чём здесь дело, надо сперва понять, какой смысл мы вкладываем в понятие обратной функции.

А смысл таков. Пусть функция $f : X \rightarrow Y$ задана, т. е. для каждого $x \in X$ мы можем указать однозначно точку $y \in Y$, являющуюся образом точки x ($y = f(x)$). Попробуем решить обратную задачу, т. е. по произвольному элементу $y \in Y$ указать точку $x \in X$, которая переводится в него функцией f , найти прообраз точки y . Если эта обратная задача разрешима, т. е. для любой точки $y \in Y$ мы можем указать единственную точку $x \in X$ такую, что $f(x) = y$, то мы тем самым построили некоторую функцию из множества Y во множество X . Она, разумеется, и будет функцией f^{-1} . Попросту говоря, обратная функция указывает для каждого элемента $y \in Y$ ту единственную точку, которая в него переходит.

Теперь ясно, что обратная функция существует далеко не всегда (это и является причиной, по которой не всегда прообраз множества можно считать его образом при обратном отображении). Например, рассмотрим функцию $f(x) = x^2$ ($f : \mathbf{R} \rightarrow \mathbf{R}$). Ясно, что взяв в качестве y число 4, мы не сможем *однозначно* установить то число, которое в него отобразилось: это может быть и 2, и -2 . Более того, если в качестве y взять число -4 ,

мы не сможем найти никакого числа, которое в него отображается: квадрат действительного числа отрицательным быть не может.

Анализ этого примера приводит к ограничениям, которые нужно наложить на функцию, если мы желаем, чтобы она имела обратную. Для строгой формулировки этих понятий нам потребуются следующие три определения.

3. Инъекция, сюръекция и биекция. Связь биекции и обратной функции

Определение 15. Функция $f : A \rightarrow B$ называется

- а) *инъекцией*, если $f(a_1) \neq f(a_2)$ для любых различных элементов множества A ,
- б) *сюръекцией*, если для любого элемента $b \in B$ найдётся $a \in A$ такой, что $f(a) = b$,
- в) *биекцией*, если f является одновременно инъекцией и сюръекцией.

Поясним эти определения. Прежде всего примеры: $f = x^2$, рассматриваемая как функция из \mathbf{R} в \mathbf{R} инъекцией не является: $f(2) = f(-2) = 4$, хотя $2 \neq -2$. Не будет она и сюръекцией: если в качестве b взять любое отрицательное число, невозможно найти a , требуемого в определении сюръекции.

Та же функция, рассмотренная как функция $[0; +\infty)$ в \mathbf{R} уже будет инъекцией: различные неотрицательные числа при возведении в квадрат не могут дать одно и то же число. Сюръекцией же эта функция по-прежнему не будет.

Если мы желаем получить сюръекцию, достаточно рассмотреть всё ту же функцию $y = x^2$, но как функцию из \mathbf{R} в $[0; +\infty)$. В самом деле, для любого неотрицательного числа a можно указать число, которое функцией f в него переводится; подойдёт число \sqrt{a} . Однако, в этом случае нет инъекции.

Если же мы рассмотрим функцию $y = x^2$ как функцию из $[0; +\infty)$ в $[0; +\infty)$ (или, например, как функцию из $[0; 1]$ в $[0; 1]$), то получим сюръекцию и инъекцию одновременно, т. е. мы получим биекцию.

Не следует думать, что три введенные понятия применимы только для вещественных функций. Рассмотрим, например, функцию, определенную на множестве всех зрителей на каком-нибудь конкретном спектакле, и ставящую в соответствие каждому такому зрителю ряд и место, прописанное в его билете. Тогда тот факт, что указанное отображение является сюръекцией, означает, что все места заняты, т. е. аншлаг. А если вдруг рассмотренная функция не будет инъекцией, возникнет конфликт: на некоторое кресло в зрительном зале будут законно претендовать несколько зрителей.

Думаю, теперь вам понятен смысл введенных понятий. Сюръекция означает, что функция может принять *все* значения из того множества, в которое она действует; не случайно в русской терминологии вместо слова «сюръекция» применяется термин *накрывающее отображение* или коротко *отображение на*. Инъективность функции означает, что точки при данном отображении не могут «склеиться», т. е. разные точки переходят в разные.

В русской терминологии вместо слова «инъекция» применяют слово «вложение», говорящее о том, что для каждой точки множества X имеется своё место во множестве Y .

О биекции же следует сказать особо. В русской терминологии её называют *взаимно-однозначным* отображением, поясняя тем самым тот факт, что она каждому элементу $x \in X$ ставит в соответствие один и только один элемент $y \in Y$ и наоборот. Наличие биекции между двумя множествами означает, что в некотором смысле в этих множествах одно и то же количество элементов. Почему в некотором смысле? Потому что для бесконечных множеств понятие количества элементов выглядит не совсем правильно; математики используют вместо этого термин «мощность множества». Мы не будем в данном курсе углубляться в этот раздел теории множеств, отметим лишь то, что если два множества конечны, то наличие между ними биекции означает именно тот факт, что элементов в них поровну.

Упражнение 19. Какие из следующих функций (рассматриваемых как функции, действующие из \mathbf{R} в \mathbf{R}) будут инъекциями, какие сюръекциями, а какие биекциями: а) $y = \operatorname{tg} x$; б) $y = \cos x$; в) $y = 2^x$; г) $y = \sqrt[5]{x}$; д) $y = x^3 - x$; е) $y = \frac{4}{x}$ при $x \neq 0$ и $y = 0$ при $x = 0$?

Упражнение 20. Пусть функции $f : X \rightarrow Y$ и $g : Y \rightarrow Z$ — биекции. Докажите, что композиция $g \circ f$ также является биекцией.

Теорема 2. Функция $f : X \rightarrow Y$ имеет обратную в том и только том случае, когда она является биекцией.

Доказательство. Необходимость. Пусть функция $f^{-1} : Y \rightarrow X$ существует. Тогда для любого элемента $y \in Y$ существует его образ во множестве X : $x = f^{-1}(y)$. По определению обратной функции $f(x) = f(f^{-1}(y)) = y$, поэтому отображение f — сюръекция. Докажем, что f — инъекция. Предположим противное, т. е. предположим, что есть два элемента $x_1 \neq x_2$ во множестве X , для которых $f(x_1) = f(x_2) = y$. Тогда $f^{-1}(y) = f^{-1}(f(x_1)) = x_1$ и $f^{-1}(y) = f^{-1}(f(x_2)) = x_2$, поэтому $x_1 = x_2$ — противоречие. Биективность функции f доказана.

Достаточность. Пусть функция f — биекция. Возьмём любую точку $y \in Y$. Так как f — сюръекция, найдётся такой элемент $x \in X$, что $f(x) = y$, а так как f — инъекция, то такой элемент единственен. Следовательно, мы имеем некоторую функцию $g : Y \rightarrow X$, ставящую в соответствие каждому элементу $y \in Y$ его прообраз во множестве X . При этом очевидно, что $f(g(y)) = y$ и так же ясно, что $g(f(x)) = x$ для любой точки $x \in X$, т. е. $g = f^{-1}$. Теорема 2 доказана.

Итак, мы установили, в частности, что функция $y = x^2$ обратной не имеет. Легко убедиться, что не имеют обратной функции ни показательная (не сюръекция), ни функция тангенса (не инъекция), ни функция синуса (ни то, ни другое). Однако мы со школы слышали, что функция

\sqrt{x} — обратная к функции x^2 , что логарифмическая функция обратна показательной, и что все основные тригонометрические функции имеют обратные (так называемые арктангенс, арксинус и т. п.). Нет ли здесь противоречия?

Противоречия на самом деле нет. Весь фокус в том, что говоря об обратных функциях к указанным, математики предполагают, что рассматриваются функции, действующие не из всей числовой прямой во всю числовую прямую, а действующие из некоторого подмножества \mathbf{R} в другое подмножество \mathbf{R} . Причём эти подмножества заранее известны, об этом имеется договорённость между математиками всего мира. Так, по умолчанию считают, что когда идёт речь, о функции, обратной к $y = x^2$, эта последняя действует из $[0; +\infty)$ в $[0; +\infty)$, а когда говорят о показательной и обратной к ней, имеют в виду, что a^x действует из \mathbf{R} в $(0; +\infty)$ (а обратная к ней логарифмическая из интервала $(0; +\infty)$ в \mathbf{R}). Что до функций, обратных к основным тригонометрическим, то мы о них поговорим подробно в следующей лекции.

А эту лекцию мы завершим весьма полезной, хоть и несложной, теоремой.

Теорема 3. Пусть $X \subset \mathbf{R}$ и $Y \subset \mathbf{R}$, функция $f : X \rightarrow Y$ имеет обратную. Тогда графики функций f и f^{-1} симметричны относительно прямой $y = x$.

Доказательство. Возьмём произвольную точку $M(x_M, y_M) \in \Gamma(f)$. Условие $M(x_M, y_M) \in \Gamma(f)$ означает, что $f(x_M) = y_M$, а это равносильно тому, что $x_M = f^{-1}(y_M)$, т. е. тому, что точка $N(y_M, x_M)$ лежит на графике функции f^{-1} . Покажем, что точка N является образом точки M при симметрии относительно прямой $y = x$, из этого будет следовать, что образ всего $\Gamma(f)$ при указанной симметрии лежит во множестве $\Gamma(f^{-1})$. Возможны два случая: 1) $x_M = y_M$, тогда точка M лежит на прямой $y = x$. Ясно, что при симметрии она переходит в себя. Осталось заметить, что в этом случае $M = N$; 2) $x_M \neq y_M$, тогда точки M и N различ-

ны; уравнение прямой MN имеет вид $y = kx + b$. Подставим в уравнение последовательно координаты обеих точек M и N , получим систему
$$\begin{cases} y_M = kx_M + b \\ x_M = ky_M + b \end{cases},$$
 решая которую, найдём, что $k = -1$, $b = x_M + y_M$. Значит, уравнение прямой MN таково: $y = -x + (x_M + y_M)$. Видно, что эта прямая параллельна прямой $y = -x$, т. е. перпендикулярна прямой $y = x$, и нам осталось убедиться, что расстояния от точек M и N до прямой $y = x$ одинаковы. Последнее почти очевидно, так как середина отрезка MN имеет координаты $\left(\frac{x_M + y_M}{2}, \frac{x_M + y_M}{2}\right)$, поэтому лежит на прямой $y = x$.

Итак, $\Gamma(f)$ при симметрии относительно прямой $y = x$ отображается на часть графика $\Gamma(f^{-1})$. Но так как функция f в свою очередь является обратной для функции f^{-1} , то верно также и то, что при той же симметрии график $\Gamma(f^{-1})$ отображается на часть графика $\Gamma(f)$. Это означает симметричность обоих графиков относительно прямой $y = x$. Теорема 3 доказана.

Упражнение 21. Рассмотрим функцию $f(x) = 6x + |x - 2| + |2x + 4|$ ($f : \mathbf{R} \rightarrow \mathbf{R}$):

- а) покажите, что функция $f(x)$ имеет обратную и постройте графики прямой и обратной функций;
- б) используя четыре арифметические операции и операцию взятия модуля, запишите аналитический вид функции f^{-1} .

ЛЕКЦИЯ IV. ТРИГОНОМЕТРИЯ

1. Единичная окружность и основные тригонометрические функции

На сегодняшней лекции мы вспомним все тригонометрические определения и факты, выведем основные формулы тригонометрии, поговорим о тригонометрических функциях и обратных к ним, посмотрим на графики этих функций. Начнём с единичной окружности.

Единичной окружностью называется окружность с центром в начале координат радиуса 1. Это, конечно, определение, но настолько очевидное, что мы не будем выделять его в качестве такового. Это же относится и к следующим нескольким понятиям, о которых сейчас пойдёт речь. Заметим также, что уравнение единичной окружности имеет вид $x^2 + y^2 = 1$.

Отметим точку с координатами $(1, 0)$ (она, очевидно, лежит на единичной окружности) и обозначим её как P_0 . Теперь рассмотрим произвольное действительное число a и поставим в соответствие этому числу некоторую точку P_a единичной окружности. Сделаем это в соответствии со следующими правилами:

- 1) Числу 0 поставим в соответствие точку P_0 .
- 2) Если $a > 0$, рассмотрим дугу единичной окружности длины a с началом в точке P_0 и отложенную против часовой стрелки. Конец этой дуги и будет точкой P_a .
- 3) Если $a < 0$, рассмотрим дугу единичной окружности длины $|a| = -a$ с началом в точке P_0 и отложенную по часовой стрелки. Конец этой дуги будет точкой P_a .

Мы получили функцию, действующую из числовой прямой в единичную окружность. Эта функция является сюръекцией, но не инъекцией, так как в каждую точку окружности отобразилось бесконечное количество точек числовой прямой. Точнее, любые две точки числовой прямой, находящиеся на расстоянии, равном длине всей единичной окружности, отобразятся в одну точку. Так как длина единичной окружности равна 2π , то это означает, что для любого числа a имеет место совпадение точек P_a и $P_{a+2\pi}$. Таким образом, $\dots = P_{a-4\pi} = P_{a-2\pi} = P_a = P_{a+2\pi} = P_{a+4\pi} = \dots$. Также ясно, что для любых чисел a и b , где $a < b$, отрезок $[a; b]$ отображается на дугу $P_a P_b$, отложенную против часовой стрелки. Если при этом $b - a < 2\pi$, то отрезок отобразится взаимно-однозначно. В частности, для любого числа a интервал $(a; a + 2\pi)$ взаимно однозначно отобразится на всю единичную окружность, исключая точку P_a . Это значит, в частности, что в точку P_a отобразятся только точки вида $a + 2\pi n$, где $n \in \mathbf{Z}$.

Весьма полезно следующее наблюдение. Точки P_a и P_{-a} , как следует из

определения, симметричны относительно оси абсцисс. Симметрия относительно начала координат эквивалентна повороту на 180° , поэтому точки P_a и $P_{a-\pi}$ (равно $P_{a+\pi}$) симметричны относительно начала координат. Наконец, симметрия относительно оси ординат получится, если последовательно осуществить симметрию относительно оси абсцисс, а затем относительно начала координат, поэтому точки P_a и $P_{\pi-a}$ симметричны относительно оси ординат. (На рис. 1 изображены все указанные точки.) Учитывая предыдущие рассуждения, получим, что точки P_a и P_b симметричны 1) относительно оси абсцисс тогда и только тогда, когда $b = -a + 2\pi n$, ($n \in \mathbf{Z}$); 2) относительно начала координат тогда и только тогда, когда $b = a - \pi + 2\pi n = a + \pi(2n - 1)$, ($n \in \mathbf{Z}$); 3) относительно оси ординат тогда и только тогда, когда $b = \pi - a + 2\pi n$, ($n \in \mathbf{Z}$).

Ещё один важный момент. Положение точки на единичной окружности может описываться не только длиной дуги P_0P_α , но и её градусной мерой, т. е. величиной центрального угла P_0OP_α . Разумеется, мы должны предполагать, что угол может быть больше 180° и даже больше 360° , а также может иметь и отрицательное число градусов. (Последнее означает, что угол откладывается от P_0 в нижнюю полуплоскость, по направлению часовой стрелки). Связь между длиной дуги и величиной центрального угла хорошо известна: центральный угол в α градусов высекает на окружности радиуса R дугу, равную $\frac{\pi\alpha}{180^\circ}R$. Так как окружность единичная, $R = 1$, поэтому углу в α градусов соответствует точка P_x , где $x = \frac{\pi\alpha}{180^\circ}$. Это значение x называют *радианной мерой* угла α . Ещё раз подчеркнём, что радианная мера угла — это просто длина соответствующей дуги единичной окружности.

Из формулы легко вывести хорошо вам известные соотношения между радианами и градусами: $0^\circ = 0$ рад., $30^\circ = \frac{\pi}{6}$ рад., $45^\circ = \frac{\pi}{4}$ рад., $60^\circ = \frac{\pi}{3}$ рад., $90^\circ = \frac{\pi}{2}$ рад., $180^\circ = \pi$ рад. и $360^\circ = 2\pi$ рад. Разумеется, для

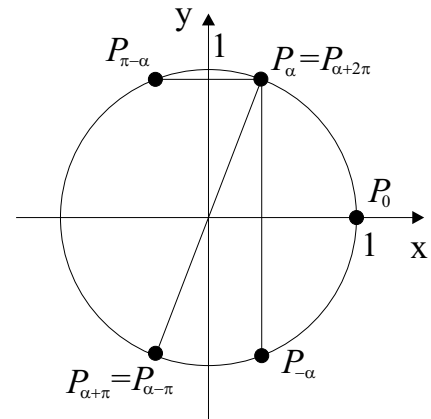


Рис. 1. Расположение точек на единичной окружности

остальных градусов тоже есть аналогичные соотношения, но приведённые углы наиболее важны с точки зрения геометрии, их называют *табличными*, знание этих углов и значений тригонометрических функций в них весьма важно.

Итак, каждому действительному числу α соответствует точка единичной окружности P_α . У этой точки (как и у всякой другой точки координатной плоскости) имеются координаты. Этими координатами и являются синус и косинус. Тангенс и котангенс определяются, как отношение этих чисел. Точнее, имеет место определение.

Определение 16. Косинусом действительного числа α (обозначение $\cos \alpha$) называется абсцисса точки P_α (рис. 2).

Синусом действительного числа α (обозначение $\sin \alpha$) называется ордината точки P_α (рис. 2).

Тангенсом действительного числа α (обозначение $\operatorname{tg} \alpha$) называется число, равное $\frac{\sin \alpha}{\cos \alpha}$, если это отношение определено.

Котангенсом действительного числа α (обозначение $\operatorname{ctg} \alpha$) называется число, равное $\frac{\cos \alpha}{\sin \alpha}$, если это отношение определено.

Не следует думать, что синус, косинус, тангенс и котангенс — единственные тригонометрические функции, существующие в природе. Например, не так давно, ещё в середине прошлого столетия, в средней школе изучались функции секанс и косеканс. В настоящий момент они практически вышли из употребления; мы даже не будем приводить здесь их определения — желающие найдут его сами в старинных учебниках или в интернете. Кроме того, тригонометрическими считаются любые функции, полученные из синуса, косинуса, тангенса и котангенса путём арифметических действий, например, функция $y = \sin 2x + \sin x$ — триго-

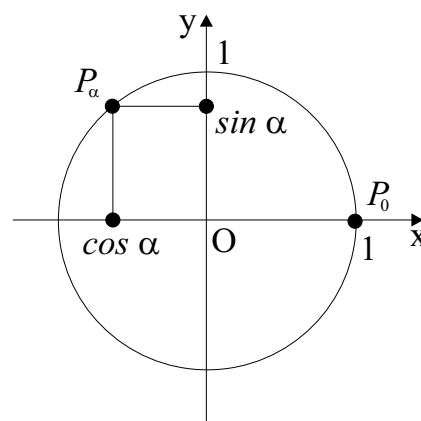


Рис. 2. Определение синуса и косинуса числа α

нометрическая. Но введённые в определении 16 тригонометрические функции считаются основными. Они так и называются: основные тригонометрические функции.

Заметим, что в средней школе понятия синуса, косинуса, тангенса и котангенса вводятся иначе, а именно, как отношение сторон прямоугольного треугольника. Да и сами эти понятия относятся не к числу, а к углу, причём изначально угол считается в градусах. Это не очень удобно, поскольку трудно распространить определение тригонометрических функций на числа, лежащие вне отрезка $\left[0; \frac{\pi}{2}\right]$. Разумнее поступать так, как это мы делаем сейчас, но это означает, что теперь нам придётся доказывать формулы, решающие прямоугольный треугольник. Например, необходимо доказать, что косинус угла прямоугольного треугольника равен отношению прилежащего катета к гипотенузе. Впрочем, это несложно.

Теорема 4. Пусть ABC — прямоугольный треугольник с прямым углом C , $\angle A = \alpha$ (в радианах). Тогда

$$\cos \alpha = \frac{AC}{AB}; \sin \alpha = \frac{BC}{AB}; \operatorname{tg} \alpha = \frac{BC}{AC}; \operatorname{ctg} \alpha = \frac{AC}{BC}.$$

Доказательство. Введём систему координат с центром в точке A . Ось абсцисс направим вдоль катета AC , ось ординат перпендикулярно этому катету и так, чтобы точка B оказалась в верхней полуплоскости. Пусть M — точка пересечения луча AB с единичной окружностью. Тогда $M = P_\alpha$, поэтому $(x_M, y_M) = (\cos \alpha; \sin \alpha)$. Точка N — проекция точки M на ось абсцисс (рис. 3). Тогда $AN = \cos \alpha$ и $MN = \sin \alpha$. Из подобия треугольников AMN и ABC находим, что $\frac{AM}{AB} = \frac{AN}{AC} = \frac{MN}{BC}$, т. е. $\frac{1}{AB} = \frac{\cos \alpha}{AC} = \frac{\sin \alpha}{BC}$. Оставляя в этом равенстве две дроби из трёх, получаем нужные соотношения. Теорема 4 доказана.

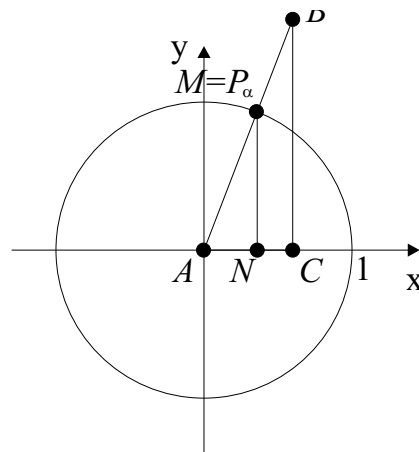


Рис. 3. К доказательству теоремы 4

Итак, зная число α , мы можем построить точку P_α (отложив на единичной окружности соответствующую дугу). Проекции этой точки на оси абсцисс и ординат будут соответственно косинусом и синусом числа α . Хотелось бы, конечно, иметь столь же наглядное геометрическое представление чисел $\operatorname{tg}\alpha$ и $\operatorname{ctg}\alpha$. И такая наглядная картинка имеется (большинство из вас её знает со школьной скамьи).

Определение 17. *Линией тангенсов* называется касательная к единичной окружности, проходящая через точку P_0 .

Линией котангенсов называется касательная к единичной окружности, проходящая через точку $P_{\frac{\pi}{2}}$.

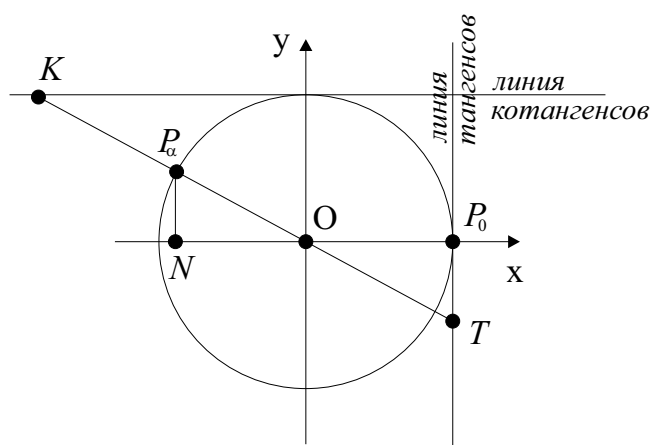


Рис. 4. Линии тангенсов и котангенсов

Если вы не совсем забыли геометрию, то сразу скажете, что линия тангенсов параллельна оси ординат и перпендикулярна оси абсцисс; аналогично линия котангенсов параллельна оси абсцисс и перпендикулярна оси ординат. Немногим сложнее записать уравнения этих прямых. Вот они. $x = 1$ — линия тангенсов, $y = 1$ — линия котангенсов.

Однако какое отношение имеют эти прямые к тангенсу и котангенсу, почему они так называются?

Проведём прямую OP_α и пусть она пересекает линию тангенсов в точке T (на рис. 4 точка P_α взята во II координатной четверти, но все рассуждения проходят и для точек, лежащих в других четвертях — убедитесь в этом сами). Пусть, кроме того, N — проекция точки P_α на ось абсцисс. Треугольники ONP_α и OTP_0 подобны по двум углам, поэтому $\frac{NP_\alpha}{ON} = \frac{P_0T}{OP_0}$. Так как $NP_\alpha = |\sin \alpha|$, $NO = |\cos \alpha|$, $OP_0 = 1$, получаем,

что $P_0T = |\operatorname{tg}\alpha|$. Кроме того, заметим, что точка T лежит ниже точки P_0 тогда и только тогда, когда P_α лежит во второй или четвёртой координатных четвертях, т. е. тогда и только тогда, когда $\operatorname{tg}\alpha$ отрицателен. Значит, ордината точки T и есть тангенс угла α . Аналогично, найдя точку пересечения прямой OP_α с линией котангенсов, получим точку K , абсцисса которой есть $\operatorname{ctg}\alpha$ (рис. 4).

2. Простейшие свойства основных тригонометрических функций

Свойства тригонометрических функций чаще всего очевидным образом следуют из их определения. Собственно говоря, вся тригонометрия может быть проиллюстрирована на единичной окружности. Начиная с этого момента и до конца лекции мы выведем большое число свойств этих функций. Чаще всего они будут записываться в виде тригонометрических формул. Мы будем нумеровать эти формулы последовательно натуральными числами, а вам рекомендуем все эти формулы знать наизусть и лучше вместе с доказательством, которое в некоторых случаях будет предшествовать формуле, а в некоторых — следовать после.

$$\sin^2 \alpha + \cos^2 \alpha = 1. \quad (1)$$

Это — так называемое основное тригонометрическое тождество. Оно немедленно следует из того, что точка $P_\alpha = (\cos \alpha; \sin \alpha)$ лежит на единичной окружности, и, следовательно, её координаты удовлетворяют уравнению этой окружности.

$$\sin(\alpha + 2\pi) = \sin \alpha; \quad \cos(\alpha + 2\pi) = \cos \alpha. \quad (2)$$

Это свойство, показывающее периодичность функций косинуса и синуса (период равен 2π) также очевидно, так как точки P_α и $P_{\alpha+2\pi}$ совпадают. Разумеется, для функций тангенс и котангенс аналогичные формулы тоже верны. Мы их не приводим, так как вскоре получим более сильные формулы. Здесь же заметим, что формулы (2) в некотором смысле неулучшаемы, т. е. какое бы число $t \in (0; 2\pi)$ мы ни взяли, равенство $\sin(\alpha + t) = \sin \alpha$

может оказаться неверным; достаточно в качестве α взять число $\frac{\pi}{2}$; убедитесь в этом самостоятельно. Аналогично неумлучшаема формула для косинуса.

$$\sin(-\alpha) = -\sin \alpha; \quad \cos(-\alpha) = \cos \alpha. \quad (3)$$

Формула (3) говорит о том, что функция синус нечётна (её график симметричен относительно начала координат), а функция косинус — чётна (график симметричен относительно оси ординат). Сама же формула следует из определения синуса и косинуса и того факта, что точки P_α и $P_{-\alpha}$ симметричны относительно оси абсцисс.

$$\begin{aligned} \sin(\alpha \pm \pi) &= -\sin \alpha; & \cos(\alpha \pm \pi) &= -\cos \alpha; \\ \sin(\pi - \alpha) &= \sin \alpha; & \cos(\pi - \alpha) &= -\cos \alpha. \end{aligned} \quad (4)$$

Эти формулы следуют из симметрии точек $P_{\alpha \pm \pi}$ и P_α относительно начала координат и из симметрии точек $P_{\pi - \alpha}$ и P_α относительно оси ординат. Есть и более общие формулы — формулы приведения. Но для их строгого вывода разумно сначала получить этот частный случай.

$$\operatorname{tg} \alpha \cdot \operatorname{ctg} \alpha = 1, \quad (\alpha \neq \frac{\pi n}{2}, \quad n \in \mathbf{Z}). \quad (5)$$

Это очевидное следствие определения тангенса и котангенса. Обратите внимание, что в формуле есть ограничение на число α . Оно здесь необходимо, поскольку (в отличие от предыдущих формул) левая и правая часть формулы определены на разных множествах; применение такой формулы без учёта этого обстоятельства может привести к ошибкам. В школьных учебниках в таком случае обычно ограничения не пишут в явном виде, просто поясняют, что формула верна только для тех α , для которых определены обе её части; мы же будем выписывать такие ограничения для наглядности и лучшей усваиваемости материала.

Выразить тангенс (котангенс) через косинус (синус) помогают следующие формулы (Они выводятся тривиально, используя основное тригонометрическое тождество — выведите их самостоятельно).

$$\operatorname{tg}^2 \alpha + 1 = \frac{1}{\cos^2 \alpha} = \frac{1}{1 - \sin^2 \alpha}; \quad \operatorname{ctg}^2 \alpha + 1 = \frac{1}{\sin^2 \alpha} = \frac{1}{1 - \cos^2 \alpha}. \quad (6)$$

Заметим, что хотя формулы (6) имеют место не для всех α , применять их можно не размышляя, так как обе части этих формул определены на одном и том же множестве.

Очевидны также следующие формулы (их можно доказать, применяя формулы (3) и (4) — сделайте это сами)

$$\begin{aligned}\operatorname{tg}(-\alpha) &= -\operatorname{tg}\alpha; & \operatorname{ctg}(-\alpha) &= -\operatorname{ctg}\alpha; \\ \operatorname{tg}(\alpha + \pi) &= \operatorname{tg}\alpha; & \operatorname{ctg}(\alpha + \pi) &= \operatorname{ctg}\alpha.\end{aligned}\tag{7}$$

Эти формулы говорят, что обе функции и тангенс и котангенс нечётные и периодические с периодом π (в отличие от синуса и косинуса, наименьший положительный период которых 2π). Доказать, что π — *наименьший* положительный период этих функций чуть сложнее. Например, для функции $\operatorname{ctg}x$ можно сделать это так: на полуинтервале $(0; \frac{\pi}{2})$ синус и косинус положительны, при этом косинус убывает (от 1 до 0), а синус возрастает (от 0 до 1). Значит, котангенс на этом промежутке, как отношение косинуса к синусу, положителен и монотонно убывает, значит, каждое своё значение принимает ровно один раз. Такие же рассуждения показывают монотонное убывание котангенса на промежутке $(\frac{\pi}{2}; \pi)$, только там котангенс отрицателен. Вывод: на промежутке $(0; \pi)$ котангенс не принимает одинаковых значений, поэтому положительного периода, меньшего π , у него не существует.

Ещё к простейшим свойствам тригонометрических функций следует отнести таблицу значений этих функций для стандартных углов (см. табл.).

Таблица

Тригонометрические функции для стандартных углов

Аргумент в градусах	0	30	45	60	90
Аргумент в радианах	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
$\sin x$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1
$\cos x$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0
$\operatorname{tg} x$	0	$1/\sqrt{3}$	1	$\sqrt{3}$	∞
$\operatorname{ctg} x$	∞	$\sqrt{3}$	1	$1/\sqrt{3}$	0

Выводятся все данные таблицы из геометрических соображений: для угла 45° рассматривается равнобедренный прямоугольный треугольник (его катеты равны), а для углов 30° и 60° — прямоугольный треугольник с этими углами (меньший его катет равен половине гипотенузы). Далее применяется теорема 4 и теорема Пифагора, находятся синус, косинус, тангенс и котангенс. Для углов 0° и 90° достаточно рассмотреть точки P_0 и $P_{\frac{\pi}{2}}$. Подробности опустим, желающие легко проведут доказательство сами или найдут его в стандартном школьном учебнике. Заметим две вещи: во-первых, в силу свойств (2) — (4) и (7) достаточно рассматривать только стандартные углы в первой координатной четверти, что и отражено в таблице. Во-вторых, в тех местах таблицы, где тангенс или котангенс не определены мы поставили символ бесконечности. Этим самым мы подчёркиваем, что хотя в самой точке тангенс (котангенс) не определён, при значениях аргумента близких к этой точке модуль тангенса (котангенса) будет неограниченно большим. В математическом анализе этот факт строго определяется через понятие предела; пока же достаточно и такого упрощённого понимания.

Часто спрашивают, как можно запомнить таблицу стандартных углов. Единой методики, понятно, нет, но вот следующее наблюдение может помочь. Если записать числа в строке синусов в виде $\frac{\sqrt{0}}{2}, \frac{\sqrt{1}}{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{3}}{2}, \frac{\sqrt{4}}{2}$, то последовательность легко запоминается. В строке косинусов та же последовательность идёт в обратном порядке (косинус в первой четверти убывает). Для строк тангенса и котангенса достаточно запомнить число $\sqrt{3}$ и тот факт, что числа в одной половине строки обратны числам в другой. Также необходимо знать, что тангенс в первой четверти возрастает, а котангенс убывает. Впрочем, как уже говорилось, способ запоминания у каждого свой.

3. Тригонометрические формулы более сложного вида

Довольно часто в заданиях по тригонометрии встречаются выражения вида $\operatorname{tg}(x + y)$, $\cos(\alpha \cdot x)$ и им подобные, выражения, в которых тригонометрическая функция берётся от аргумента, представляющего собой сумму или произведение. К сожалению, многие выпускники средней школы делают ошибки в их трактовке. Например, некоторые студенты считают, что множитель константу можно вынести за знак функции (т. е. по их мнению, например, $\sin 3x = 3 \sin x$), другие почему-то заменяют косинус разности на разность косинусов и т. д. Когда же выясняется (чаще всего им об этом объявляет преподаватель), что так поступать нельзя, студенты попросту не представляют, как нужно действовать. Между тем в тригонометрии имеется огромное число формул, позволяющих преобразовать такие выражения. Хотя многие из этих формул известны со школы, мы приведём как сами формулы, так и их вывод. Выводы полезно знать, поскольку самих формул много и механически все запомнить тяжело. Формулы тесно связаны друг с другом, одна вытекает из другой, и зная эту связь несложно восстановить пробелы. А связь эта хорошо просматривается в выводах формул.

Начнём с формулы косинуса суммы двух чисел, единственной формулы, для которой нам понадобятся вещи не совсем тригонометрические: теорема косинусов и формула расстояния между двумя точками на координатной плоскости. Поэтому вывод формулы косинуса суммы сложнее прочих, и эту формулу лучше помнить наизусть.

$$\cos(\alpha + \beta) = \cos \alpha \cdot \cos \beta - \sin \alpha \cdot \sin \beta. \quad (8)$$

Вывод формулы проводится в несколько этапов.

Этап 1. Пусть $\alpha, \beta \in [0; \pi]$. Рассмотрим точки P_α и $P_{-\beta}$. Они лежат в верхней и нижней полуплоскостях соответственно. Рассмотрим треугольник $P_\alpha OP_{-\beta}$. Если $\alpha + \beta < \pi$, то $\angle P_\alpha OP_{-\beta} = \angle P_\alpha OP_0 + \angle P_0 OP_{-\beta} = \alpha + \beta$, если же $\alpha + \beta > \pi$, то $\angle P_\alpha OP_{-\beta} = 2\pi - (\alpha + \beta)$. В обоих случаях получим, применив к треугольнику $P_\alpha OP_{-\beta}$ теорему косинусов, что $P_\alpha P_{-\beta}^2 = OP_\alpha^2 + OP_{-\beta}^2 - 2OP_\alpha OP_{-\beta} \cos(\alpha + \beta)$. Но $OP_\alpha = OP_{-\beta} = 1$, как

радиусы единичной окружности, и видим, что $P_\alpha P_{-\beta}^2 = 2 - 2 \cos(\alpha + \beta)$. Для нахождения $P_\alpha P_{-\beta}^2$ можно воспользоваться формулой расстояния между точками координатной плоскости, ведь координаты точек P_α и $P_{-\beta}$ известны: $P_\alpha(\cos \alpha, \sin \alpha)$, $P_{-\beta}(\cos \beta, -\sin \beta)$. Имеем

$$2 - 2 \cos(\alpha + \beta) = P_\alpha P_{-\beta}^2 = (\cos \alpha - \cos \beta)^2 + (\sin \alpha + \sin \beta)^2.$$

Раскрыв скобки в правой части и дважды применив основное тригонометрическое тождество, получим

$$2 - 2 \cos(\alpha + \beta) = 2 - 2 \cos \alpha \cos \beta + 2 \sin \alpha \sin \beta,$$

откуда немедленно получается доказываемая формула. Этап 1 завершён.

Этап 2. Пусть $\alpha \in [\pi; 2\pi]$, $\beta \in [0; \pi]$ (или наоборот). Тогда $\alpha = \pi + \varphi$, где $\varphi \in [0; \pi]$ и, применив формулы (4) и результат первого этапа, имеем

$$\begin{aligned} \cos(\alpha + \beta) &= \cos(\pi + \varphi + \beta) = -\cos(\varphi + \beta) = -(\cos \varphi \cos \beta - \sin \varphi \sin \beta) = \\ &= \cos(\pi - \varphi) \cos \beta - \sin(\pi - \varphi) \sin \beta = \cos \alpha \cos \beta - \sin \alpha \sin \beta. \end{aligned}$$

Этап 2 завершён.

Этап 3. Пусть α, β — произвольные числа. Тогда $\alpha = 2\pi k + \varphi$, $\beta = 2\pi l + \psi$, где $\varphi, \psi \in [0; 2\pi]$, $k, l \in \mathbf{Z}$ и, применив периодичность функций синус и косинус (формулы (2)) и результат этапа 2, получаем

$$\begin{aligned} \cos(\alpha + \beta) &= \cos(2\pi k + \varphi + 2\pi l + \psi) = \cos(\varphi + \psi) = \\ &= \cos \varphi \cos \psi - \sin \varphi \sin \psi = \\ &= \cos(2\pi k + \varphi) \cos(2\pi l + \psi) - \sin(2\pi k + \varphi) \sin(2\pi l + \psi) = \\ &= \cos \alpha \cos \beta - \sin \alpha \sin \beta. \end{aligned}$$

Этап 3 завершён, формула доказана.

Теперь формулы пойдут обильным потоком. Во-первых, положив в формуле (8) $-\beta$ вместо β получим формулу косинуса разности.

$$\cos(\alpha - \beta) = \cos \alpha \cdot \cos \beta + \sin \alpha \cdot \sin \beta. \quad (9)$$

Далее, положив в формуле (9) $\beta = \pi/2$, и вспомнив про чётность косинуса, найдём, что $\cos\left(\frac{\pi}{2} - \alpha\right) = \sin \alpha$ (мы не будем специально выделять эту и следующую формулу, ибо они — всего лишь две из бесконечной серии *формул приведения*, о которых пойдёт речь позже). В последнем равенстве заменим α на $\frac{\pi}{2} - \alpha$ — получим, что $\sin\left(\frac{\pi}{2} - \alpha\right) = \cos \alpha$. Применим полученные два равенства для вывода синуса суммы и синуса разности: $\sin(\alpha \pm \beta) = \cos\left(\frac{\pi}{2} - (\alpha \pm \beta)\right) = \cos\left(\left(\frac{\pi}{2} - \alpha\right) \mp \beta\right) = \cos\left(\frac{\pi}{2} - \alpha\right) \cos \beta \pm \sin\left(\frac{\pi}{2} - \alpha\right) \sin \beta = \sin \alpha \cos \beta \pm \sin \beta \cos \alpha$. Мы получили следующие две формулы.

$$\sin(\alpha + \beta) = \sin \alpha \cdot \cos \beta + \cos \alpha \cdot \sin \beta. \quad (10)$$

$$\sin(\alpha - \beta) = \sin \alpha \cdot \cos \beta - \cos \alpha \cdot \sin \beta. \quad (11)$$

Воспользуемся формулами (8) — (11) для вывода формулы суммы (разности) тангенсов: $\operatorname{tg}(\alpha \pm \beta) = \frac{\sin(\alpha \pm \beta)}{\cos(\alpha \pm \beta)} = \frac{\sin \alpha \cdot \cos \beta \pm \cos \alpha \cdot \sin \beta}{\cos \alpha \cdot \cos \beta \mp \sin \alpha \cdot \sin \beta}$. Теперь поделим числитель и знаменатель на произведение $\cos \alpha \cos \beta$. Разумеется, такое деление допустимо только если оба косинуса не равны нулю; именно этот случай и описывается формулой. А формулы, как легко заметить, получаются такие:

$$\operatorname{tg}(\alpha + \beta) = \frac{\operatorname{tg} \alpha + \operatorname{tg} \beta}{1 - \operatorname{tg} \alpha \cdot \operatorname{tg} \beta}, \quad (\alpha, \beta, (\alpha + \beta) \neq \pi/2 + \pi k, k \in \mathbf{Z}). \quad (12)$$

$$\operatorname{tg}(\alpha - \beta) = \frac{\operatorname{tg} \alpha - \operatorname{tg} \beta}{1 + \operatorname{tg} \alpha \cdot \operatorname{tg} \beta}, \quad (\alpha, \beta, (\alpha - \beta) \neq \pi/2 + \pi k, k \in \mathbf{Z}). \quad (13)$$

Обратите внимание на область применимости данных формул. Левая и правая часть должны быть определены обе. И если из того, что правая часть равенства (12) (или (13)) определена, следует, что определена и его левая часть, то обратное неверно. Поэтому если формулу (12) применить, например, к выражению $\operatorname{tg}(x + \pi/4)$ при решении тригонометрического уравнения, есть возможность потерять корень типа $x = 5\pi/2$. Так что будьте осторожны, работая с этими формулами.

Совершенно аналогично выводятся формулы для котангенсов. Их вывод оставляю вам для самостоятельной работы.

$$\operatorname{ctg}(\alpha + \beta) = \frac{\operatorname{ctg} \alpha \cdot \operatorname{ctg} \beta - 1}{\operatorname{ctg} \alpha + \operatorname{ctg} \beta}, \quad (\alpha, \beta, (\alpha + \beta) \neq \pi k, k \in \mathbf{Z}). \quad (14)$$

$$\operatorname{ctg}(\alpha - \beta) = \frac{\operatorname{ctg}\alpha \cdot \operatorname{ctg}\beta + 1}{\operatorname{ctg}\beta - \operatorname{ctg}\alpha}, \quad (\alpha, \beta, (\alpha - \beta) \neq \pi k, k \in \mathbf{Z}). \quad (15)$$

Упражнение 22. Докажите формулы (14) и (15).

Далее, в формулах суммы можно положить $\alpha = \beta$. Получим формулы двойного угла.

$$\cos 2\alpha = \cos^2 \alpha - \sin^2 \alpha. \quad (16)$$

$$\sin 2\alpha = 2 \sin \alpha \cdot \cos \alpha. \quad (17)$$

$$\operatorname{tg} 2\alpha = \frac{2 \operatorname{tg} \alpha}{1 - \operatorname{tg}^2 \alpha}, \quad (\alpha \neq \frac{\pi}{4} + \frac{\pi}{2}k, \alpha \neq \frac{\pi}{2}k, k \in \mathbf{Z}). \quad (18)$$

$$\operatorname{ctg} 2\alpha = \frac{\operatorname{ctg}^2 \alpha - 1}{2 \operatorname{ctg} \alpha}, \quad (\alpha \neq \frac{\pi}{2}k, k \in \mathbf{Z}). \quad (19)$$

Часто бывает полезным формулу (16) иметь в преобразованном виде, получающимся при замене с помощью основного тригонометрического тождества косинуса на синус или наоборот.

$$\cos 2\alpha = 2 \cos^2 \alpha - 1. \quad (20)$$

$$\cos 2\alpha = 1 - 2 \sin^2 \alpha. \quad (21)$$

Сейчас легко получить формулы тройного угла: надо в формулах суммы положить $\beta = 2\alpha$ и применить формулы (16) — (19). Затем, используя полученные формулы, можно будет получить тригонометрические функции учетверённого угла, далее упятерённого и так до бесконечности. Не будем этим заниматься, но формулы синуса и косинуса тройного угла всё же приведём.

$$\cos 3\alpha = 4 \cos^3 \alpha - 3 \cos \alpha. \quad (22)$$

$$\sin 3\alpha = 3 \sin \alpha - 4 \sin^3 \alpha. \quad (23)$$

Упражнение 23. Докажите формулы (22) и (23).

Теперь заменим в формулах (20) и (21) α на $\alpha/2$. После очевидных преобразований получим формулы половинного аргумента.

$$\cos^2 \frac{\alpha}{2} = \frac{1 + \cos \alpha}{2}. \quad (24)$$

$$\sin^2 \frac{\alpha}{2} = \frac{1 - \cos \alpha}{2}. \quad (25)$$

Обратите внимание, что часто встречающаяся запись этих формул через квадратный корень (например, формулу (24) иногда записывают в виде $\cos \frac{\alpha}{2} = \pm \sqrt{\frac{1 + \cos \alpha}{2}}$, а иногда и вовсе опускают знак \pm перед радикалом) не вполне корректна: при конкретном значении α косинус и синус половинного угла определены однозначно, они могут быть как положительными, так и отрицательными, но не одновременно теми и другими. Это же замечание относится к следующим двум формулам, получающимся делением друг на друга формул (24) и (25).

$$\operatorname{tg}^2 \frac{\alpha}{2} = \frac{1 - \cos \alpha}{1 + \cos \alpha}. \quad (26)$$

$$\operatorname{ctg}^2 \frac{\alpha}{2} = \frac{1 + \cos \alpha}{1 - \cos \alpha}. \quad (27)$$

В этих формулах нет необходимости указывать ограничения на α , поскольку обе части каждой формулы определены на одних и тех же значениях α .

Наконец, следует отметить формулы, выражающие все основные тригонометрические функции через тангенс половинного угла. Формула для тангенса получается из формулы (18) заменой α на $\alpha/2$; перевернув получившуюся дробь, получим формулу для котангенса. Формула для косинуса получается тривиальным преобразованием формулы (26); наконец, формулу для синуса можно получить умножив формулу для тангенса на формулу для косинуса. Получатся следующие формулы, весьма полезные в математическом анализе (например, при интегрировании тригонометрических функций).

$$\operatorname{tg} \alpha = \frac{2 \operatorname{tg}(\alpha/2)}{1 - \operatorname{tg}^2(\alpha/2)}. \quad (28)$$

$$\operatorname{ctg} \alpha = \frac{1 - \operatorname{tg}^2(\alpha/2)}{2 \operatorname{tg}(\alpha/2)}. \quad (29)$$

$$\cos \alpha = \frac{1 - \operatorname{tg}^2(\alpha/2)}{1 + \operatorname{tg}^2(\alpha/2)}. \quad (30)$$

$$\sin \alpha = \frac{2 \operatorname{tg}(\alpha/2)}{1 + \operatorname{tg}^2(\alpha/2)}. \quad (31)$$

Применять эти формулы следует с большой осторожностью: их правая часть не определена, если не определён $\operatorname{tg}(\alpha/2)$, т. е. не определена при $\alpha = \pi + 2\pi k$, ($k \in \mathbf{Z}$), в то время, как левая часть в этих точках определена (исключая формулу для котангенса). Не учёт этого факта может вполне (как уже отмечали сегодня) привести к потере корней уравнения или неравенства.

Важнейшим приложением формул суммы и разности являются формулы приведения. Собственно говоря, формулы приведения — это и есть формулы суммы и разности для случая, когда $\beta = k\pi/2$ для некоторого натурального k (вы, конечно, понимаете, что это означает то, что точка P_β лежит на пересечении единичной окружности и одной из координатных осей). Каждая из этих формул легко получается прямым раскрытием скобок по формулам (8) — (15). Однако существует общее правило их получения, которое мы приведем без вывода.

Правило (формулы приведения). Пусть f — основная тригонометрическая функция, $k \in \mathbf{Z}$. Тогда выражение $f\left(x + \frac{\pi}{2}k\right)$ преобразуется последовательным выполнением двух пунктов:

- 1) Если число k чётно, то функция f остаётся без изменений, если же k нечётно, то она заменяется на кофункцию, т. е. синус на косинус, косинус на синус, тангенс на котангенс и котангенс на тангенс. Аргументом функции во всех случаях остаётся x (выражение $\frac{\pi}{2}k$ отбрасывается).
- 2) Знак перед функцией может оказаться как плюс, так и минус. Этот знак совпадает со знаком выражения $f\left(x + \frac{\pi}{2}k\right)$ при x , лежащем в 1-й координатной четверти (т. е. при малых положительных x).

Покажем на конкретном примере как действует это правило. Пусть имеется выражение $\cos\left(3\alpha + \frac{5\pi}{2}\right)$. Так как число 5 нечётно, то по первому пункту правила мы должны заменить косинус на синус, а аргумент — просто на 3α . Определим знак, для чего будем полагать, что 3α лежит в первой четверти. Тогда ясно, что точка $3\alpha + \frac{5\pi}{2}$ попадает во вторую четверть, косинус там отрицателен, и значит мы должны ставить знак минус.

Окончательно имеем $\cos\left(3\alpha + \frac{5\pi}{2}\right) = -\sin 3\alpha$.

Продолжим вывод формул тригонометрии. Сложив друг с другом формулы (8) и (9) (имеется в виду, что мы складываем отдельно левые и отдельно правые части формул), мы получим формулу произведения косинусов, а вычитая их — формулу произведения синусов.

$$\cos \alpha \cdot \cos \beta = \frac{\cos(\alpha + \beta) + \cos(\alpha - \beta)}{2}. \quad (32)$$

$$\sin \alpha \cdot \sin \beta = \frac{\cos(\alpha - \beta) - \cos(\alpha + \beta)}{2}. \quad (33)$$

Точно также сложим формулы (10) и (11). Получим произведение синуса на косинус.

$$\sin \alpha \cdot \cos \beta = \frac{\sin(\alpha + \beta) + \sin(\alpha - \beta)}{2}. \quad (34)$$

Вычитать формулы (10) и (11) бесполезно — получим ту же формулу произведения синуса на косинус. Формулы произведения тангенсов и котангенсов интереса не представляют, так как являются отношениями произведений синусов и произведений косинусов.

Положим в формулах (32) — (34) $\alpha = \frac{x+y}{2}$, $\beta = \frac{x-y}{2}$. Тогда $\alpha + \beta = x$, $\alpha - \beta = y$, и мы получаем формулы суммы тригонометрических функций.

$$\cos x + \cos y = 2 \cos \frac{x+y}{2} \cdot \cos \frac{x-y}{2}. \quad (35)$$

$$\cos x - \cos y = 2 \sin \frac{y-x}{2} \cdot \sin \frac{x+y}{2}. \quad (36)$$

$$\sin x + \sin y = 2 \sin \frac{x+y}{2} \cdot \cos \frac{x-y}{2}. \quad (37)$$

Обратите внимание на то, что если в этих формулах положить $y = 0$, то получатся выведенные ранее формулы половинного аргумента. Ещё бывает разность синусов. Эта формула немедленно получается из формулы (37) заменой y на $-y$.

$$\sin x - \sin y = 2 \sin \frac{x-y}{2} \cdot \cos \frac{x+y}{2}. \quad (38)$$

Формулы для суммы синуса и косинуса специальной нет. Она и не нужна, так как в силу формул приведения такая сумма может быть легко сведена

как к сумме синусов так и к сумме косинусов. Например, $\sin x + \cos x = \cos\left(\frac{\pi}{2} - x\right) + \cos x = 2 \cos \frac{\frac{\pi}{2} - x + x}{2} \cos \frac{\frac{\pi}{2} - x - x}{2} = \sqrt{2} \cos\left(\frac{\pi}{4} - x\right)$ или $\sin x + \cos x = \sin x + \sin\left(\frac{\pi}{2} + x\right) = 2 \sin \frac{x + \frac{\pi}{2} + x}{2} \cos \frac{\frac{x - \pi}{2} - x}{2} = \sqrt{2} \sin\left(\frac{\pi}{4} + x\right)$. Эти конкретные формулы встречаются часто; выделим их нумерацией.

$$\sin x + \cos x = \sqrt{2} \cos\left(\frac{\pi}{4} - x\right) = \sqrt{2} \sin\left(\frac{\pi}{4} + x\right). \quad (39)$$

$$\sin x - \cos x = -\sqrt{2} \cos\left(\frac{\pi}{4} + x\right) = \sqrt{2} \sin\left(x - \frac{\pi}{4}\right). \quad (40)$$

Ещё имеются формулы суммы и разности тангенсов и котангенсов, хотя они используются реже. Мы приведём формулы без доказательства — докажете их сами.

$$\operatorname{tg} x + \operatorname{tg} y = \frac{\sin(x + y)}{\cos x \cos y}. \quad (41)$$

$$\operatorname{tg} x - \operatorname{tg} y = \frac{\sin(x - y)}{\cos x \cos y}. \quad (42)$$

$$\operatorname{ctg} x + \operatorname{ctg} y = \frac{\sin(x + y)}{\sin x \sin y}. \quad (43)$$

$$\operatorname{ctg} x - \operatorname{ctg} y = \frac{\sin(y - x)}{\sin x \sin y}. \quad (44)$$

Упражнение 24. Докажите формулы (41) — (44).

На этом список основных тригонометрических формул стоит завершить. Обратите внимание насколько эти формулы взаимосвязаны; любую из них, если вдруг подзабылась, легко восстановить используя другие формулы. Если примерно представлять общий вид формулы, то детали (знаки и константы) можно восстановить, посмотрев частные случаи. Пользуйтесь этим при работе с тригонометрией.

4. Функции, обратные основным тригонометрическим, и их графики

Прежде всего напомним понятие арксинуса, арккосинуса, арктангенса и арккотангенса числа — в средней школе этим понятиям уделяется меньше

времени, чем нужно, поэтому многие могли эти понятия подзабыть.

Определение 18: а) *Арксинусом* числа x (обозначается $\arcsin x$) называют число, лежащее в отрезке $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$, синус которого равен x ;

б) *Арккосинусом* числа x ($\arccos x$) называют число, лежащее в отрезке $[0; \pi]$, косинус которого равен x ;

в) *Арктангенсом* числа x ($\arctg x$) называют число, лежащее в интервале $\left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$, тангенс которого равен x ;

г) *Арккотангенсом* числа x ($\operatorname{arcctg} x$) называют число, лежащее в интервале $(0; \pi)$, котангенс которого равен x .

В приведённом выше определении следует обратить внимание на два момента:

1) Все определённые понятия — это не углы, не функции, это именно *числа*, элементы \mathbf{R} . Пока что, с одной стороны, не идёт речи ни о каких обратных функциях, а с другой стороны, мы имеем дело не с прямоугольным треугольником и его углами, а со всей числовой прямой.

2) Для каждого из чисел $\arcsin x$, $\arccos x$, $\arctg x$, $\operatorname{arcctg} x$ прямо в определении указано некоторое подмножество числовой прямой, которому это число будет заведомо принадлежать. Это подмножество своё для каждого числа и легко понять, из каких соображений оно выбрано.

Понятие арксинуса введено для того, чтобы решать уравнения вида $\sin x = a$ (где $|a| \leq 1$), т. е. число $\arcsin a$ должно быть одним из решений этого уравнения. Каким именно? Заметим, что если переменная x пробегает отрезок $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$, то функция $\sin x$ принимает все свои значения и каждое ровно по одному разу. Кроме того, указанный отрезок содержит наиболее важную с практической точки зрения часть числовой прямой — промежуток от 0 до $\pi/2$ радиан, соответствующую углам от 0° до 90° . Других отрезков (и вообще связанных промежутков), удовлетворяющих этим двум условиям, просто нет. Таким образом, естественно потребовать, чтобы число $\arcsin a$ лежало именно в указанном отрезке. Рассуждения в случае $\arccos x$, $\arctg x$, $\operatorname{arcctg} x$ аналогичны.

Если определение 18 записать формально, получим следующую запись,

которую мы рекомендуем хорошо понять и запомнить:

$$y = \arcsin a \Leftrightarrow \begin{cases} a = \sin y \\ -\frac{\pi}{2} \leq y \leq \frac{\pi}{2} \end{cases}; \quad y = \arccos a \Leftrightarrow \begin{cases} a = \cos y \\ 0 \leq y \leq \pi \end{cases};$$

$$y = \operatorname{arctg} a \Leftrightarrow \begin{cases} a = \operatorname{tg} y \\ -\frac{\pi}{2} < y < \frac{\pi}{2} \end{cases}; \quad y = \operatorname{arcctg} a \Leftrightarrow \begin{cases} a = \operatorname{ctg} y \\ 0 < y < \pi \end{cases}.$$

Теперь становится окончательно понятно, почему, например, $\arcsin(\sin 10) \neq 10$, хотя $\sin(\arcsin x) = x$ для всех x , для которых определено понятие арксинуса. В самом деле, тождество после слова «хотя» совершенно очевидно из определения: если $\arcsin x = b$, то $\sin b = x$ и $\sin(\arcsin x) = \sin b = x$. Чтобы найти, чему равен $\arcsin(\sin 10)$, обозначим это выражение буквой x и (опять-таки из определения арксинуса)

получим систему $\begin{cases} \sin x = \sin 10 \\ -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \end{cases}$. Ясно, что $x = 10$ не удовлетво-

ряет второму уравнению системы, так что $\arcsin(\sin 10) \neq 10$. Решается же система либо графически (на единичной окружности), либо первое уравнение системы преобразуется по формуле разности синусов в произведение и решается. Получается, что корни первого уравнения имеют вид либо $10 + 2\pi k$, $k \in \mathbf{Z}$, либо $\pi - 10 + 2\pi k$, $k \in \mathbf{Z}$. Далее убеждаемся, что среди корней первого вида нет таких, которые бы лежали на отрезке $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$, а среди корней второго вида такой есть, причём единственный — число $3\pi - 10$. Итак, $\arcsin(\sin 10) = 3\pi - 10$.

Способ рассуждения, применённый здесь, стандартен при решении задач такого типа. Выражение, содержащее приставку «арк» обозначается за неизвестное, записывается система, расшифровывающая эту приставку, эта система решается и выписывается ответ. Применим этот метод для доказательства одной известной формулы, а именно докажем, что $\sin(\arccos x) = \sqrt{1 - x^2}$ при всех $x \in [-1; 1]$.

Пусть $y = \arccos x$. Тогда $x = \cos y$ и $0 \leq y \leq \pi$. Значит, $\sin y \geq 0$. Из основного тригонометрического тождества имеем

$$\sin y = \sqrt{1 - \cos^2 y} = \sqrt{1 - x^2},$$

что и требовалось доказать.

Упражнение 25. Найдите, чему равны числа: а) $\operatorname{arctg}(\operatorname{tg}(-3))$;
б) $\cos(\arccos 0,8 + \arccos 0,6)$.

Упражнение 26. а) Докажите, что для всех чисел x из отрезка $[-1; 1]$ справедлива формула $\arcsin x + \arccos x = \frac{\pi}{2}$. б) Выведите и докажите аналогичную формулу для суммы арктангенса и арккотангенса одного и того же аргумента.

Упражнение 27. Постройте график функции $y = \arccos(\cos x)$.

Ну вот мы и дошли до того момента, когда можно рассматривать функции $y = \arccos x$, $y = \arcsin x$, $y = \operatorname{arctg} x$ и $y = \operatorname{arcctg} x$. Все эти функции считаются определёнными на максимально возможном подмножестве числовой прямой, т. е. функции $y = \operatorname{arctg} x$ и $y = \operatorname{arcctg} x$ определены на \mathbf{R} (так как и тангенс, и котангенс могут принимать любые действительные значения), а функции $y = \arccos x$ и $y = \arcsin x$ определены на отрезке $[-1; 1]$ (синус и косинус не превосходят по модулю число 1). Ясно, что областями значения этих четырёх функций будут именно те промежутки, которые указаны в определении 18, т. е. $\arcsin x : [-1; 1] \rightarrow \left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$ и т. д. Заметим, что так определённые функции являются биекциями, значит, по теореме 2 имеют обратные. Эти обратные функции, как видно из определения, являются функциями $\sin x$, $\cos x$, $\operatorname{tg} x$ и $\operatorname{ctg} x$ суженными соответственно на множества $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$, $[0; \pi]$, $\left(-\frac{\pi}{2}; \frac{\pi}{2}\right)$ и $(0; \pi)$, а тогда выходит, что функция $y = \operatorname{arcctg} x$ является обратной к функции $y = \operatorname{ctg} x$, суженной на интервал $(0; \pi)$ (аналогично обстоит дело с тремя другими функциями). Для краткости говорят просто, что функция $\operatorname{arcctg} x$ является обратной к котангенсу (опуская тот факт, что функция рассматривается не на всей числовой прямой) — формально возникает противоречие (ведь котангенс не есть биекция), по сути же это просто договорённость.

Осталось понять, как же выглядят графики всех четырёх функций, об-

ратных к основным тригонометрическим. В этом нам поможет теорема 3. Изобразим графики функций, обратных основным тригонометрическим, используя симметрию относительно прямой $y = x$. Вот как будет выглядеть график функции $y = \arcsin x$ (рис. 5). Здесь график функции $y = \sin x$ изображён пунктирной линией; жирной пунктирной линией изображено его сужение на отрезок $\left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$ — именно к этому сужению и строится обратная функция, являющаяся арксинусом. График функции $y = \arcsin x$ — сплошная жирная линия.

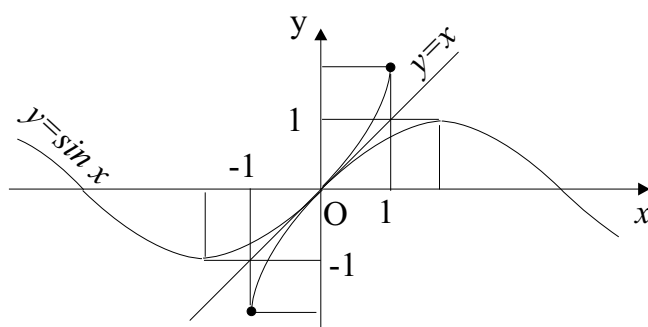


Рис. 5. График \arcsin

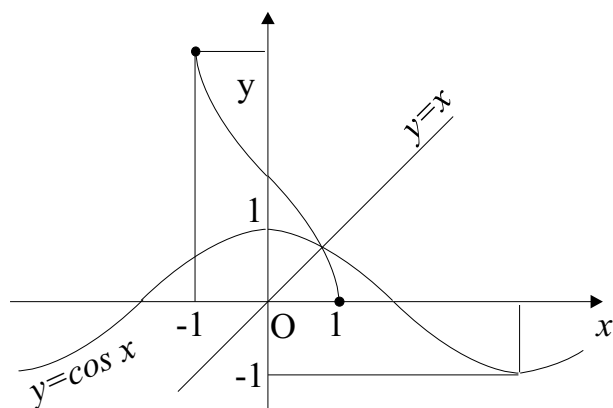


Рис. 6. График \arccos

На рис. 6 изображён график функции $y = \arccos x$. Смысл пунктирных и сплошных линий разной толщины тот же, что и при построении графика $y = \arcsin x$. Графики функций $y = \arctg x$ и $y = \text{arcctg } x$ вы можете видеть на рис. 7.

В отличие от первых двух графиков, графики функций $y = \text{tg } x$ и $y = \text{ctg } x$ изображены частично: только те их ветви, которые отображаются (относительно прямой $y = x$) при построении обратной функции. На графиках хорошо видны их горизонтальные асимптоты: у функции $y = \arctg x$ это пря-

мые $y = \pm \frac{\pi}{2}$, у функции $y = \text{arcctg } x$ асимптотами являются прямые $y = 0$ и $y = \pi$. Наличие асимптот — одно из основных отличий, например, функции $y = \arctg x$ от функции $y = \sqrt[3]{x}$ (графики этих функций кажутся похожими при построении) — последняя функция асимптот не имеет, и возрастает (в отличие от арктангенса) неограничено.

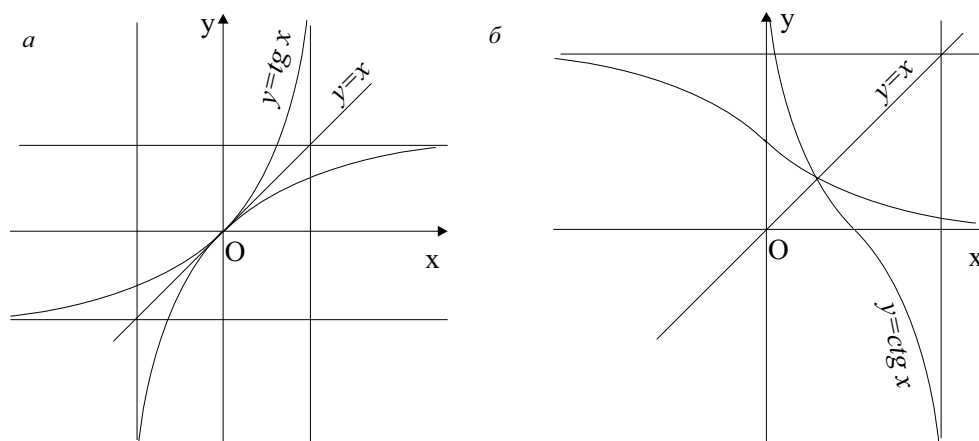


Рис. 7. Графики арктангенса (а) и арккотангенса (б)

Упражнение 28. Изобразите схематически графики функций:

- а) $y = \arccos \frac{x-3}{4}$; б) $y = 2 \operatorname{arccotg} |x-1|$; в) $y = \arcsin x + x + \arccos x$.

ЛЕКЦИЯ V. РАВНОСИЛЬНЫЕ ПЕРЕХОДЫ В УРАВНЕНИЯХ (НЕРАВЕНСТВАХ)

1. Функциональный подход к уравнениям и неравенствам

В отличие от «школьного» понимания уравнения, как равенства, содержащего переменную, мы будем считать (как это и принято среди математиков), что всякое уравнение представляет из себя равенство значений двух функций $f(x) = g(x)$. Аналогично, всякое неравенство для нас — это соотношение между функциями, стоящими в левой и правой части неравенства. Такой подход упрощает решение уравнения (неравенства), поскольку делает возможным применение хорошо развитой математической теории — математического анализа, основы которого изучаются в 10 — 11 классах средней школы. Различие со «школьным» взглядом на уравнения (неравенства) хорошо иллюстрируется следующим примером.

Пусть дано уравнение: $x^{-x} = 4$. С точки зрения школьника, его решением является всякое число (действительное), обращающее уравнение в числовое равенство, поэтому число $x_0 = -2$ — одно из решений (легко видеть, что $(-2)^{-(-2)} = 4$). С точки зрения математика речь идёт о равенстве

двух функций $f(x) = x^{-x}$ и $g(x) = 4$. Однако функция x^{-x} определяется как композиция функций e^x и $\ln x$, а именно, $x^{-x} = e^{-x \cdot \ln x}$, и поэтому она не определена при отрицательных значениях x . Значит, число -2 корнем уравнения не является.

Не будем обсуждать, какой из этих двух подходов более правильный и почему, примем тот, который более удобен, и в настоящий момент времени считается общепризнанным. Это — функциональный подход. В старших классах средней школы даётся именно он: понятие области допустимых значений уравнения (неравенства) по сути есть понятие области определения функций, в это уравнение (неравенство) входящих. Решение, например, неравенства $2^x > 4$ обосновывается через функции, а именно: функция $y = 2^x$ монотонно возрастает, определена на всей числовой прямой и принимает значение 4 в точке $x = 2$, поэтому решением данного неравенства будут все действительные числа, большие числа 2.

Итак, рассмотрим уравнение в самом общем виде $f(x) = g(x)$. Для простоты рассуждений будем полагать, что и f , и g — числовые функции, действующие из некоторых подмножеств числовой прямой в числовую прямую (общий случай аналогичен). Если в качестве x рассмотреть любое действительное число, то возможны три случая.

1) Точка x не входит в область определения по крайней мере одной из функций f или g (пусть, например, f). В этом случае не определено число $f(x)$, говорить о равенстве этого неопределённого числа чему бы то ни было абсурдно. В этом случае говорят, что число x не входит в область допустимых значений уравнения.

2) Точка x входит в область определения каждой из функций f и g и при этом значения функций в этой точке совпадают. Тогда говорят, что точка x является *корнем*, или *решением* уравнения $f(x) = g(x)$.

3) Точка x входит в область определения каждой из функций f и g , но значения функций в этой точке различны. Такие x не имеют специального названия.

Таким образом, с каждым уравнением $f(x) = g(x)$ связано два множества: множество всех точек, которые являются корнями уравнения — это

множество решений данного уравнения — и область допустимых значений уравнения. Легко видеть, что первое из них является подмножеством второго, и что второе есть *пересечение областей определения* функций f и g .

Решить уравнение — это означает найти множество его решений. Более точно, это множество нужно не только указать в явном виде, но и обосновать, что оно является искомым, т. е. обосновать, что всякий элемент этого множества является решением уравнения и наоборот. В школе эту же мысль выражают фразой «чтобы решить уравнение, необходимо указать все его корни и доказать, что других корней нет».

С неравенствами дело обстоит аналогично. Всякое неравенство может быть записано одним из четырёх способов: $f(x) < g(x)$, $f(x) \leq g(x)$, $f(x) > g(x)$ или $f(x) \geq g(x)$ для некоторых функций (по-прежнему полагаем их числовыми) f и g . Пересечение областей определения функций f и g называется областью допустимых значений неравенства, множество тех значений x из области допустимых значений, которые при подстановке в неравенство делают его верным, называется множеством решений неравенства (а сами такие x — решениями, или, точнее, *частными решениями* неравенства); решить неравенство — это значит найти его множество решений и доказать, что найденное множество искомое.

Если уравнений (неравенств) несколько, то они могут образовывать системы или совокупности, а эти системы и совокупности сами могут входить в более сложные системы и совокупности — в дальнейшем такое будет наблюдаться неоднократно. Как уже упоминалось при изучении множеств (см. лекцию 1), множество решений совокупности — это *объединение* множеств решений объектов, входящих в совокупность, а множество решений системы — *пересечение* таких множеств.

2. Равносильные и неравносильные уравнения (неравенства)

Определение 19. Два уравнения (два неравенства, уравнение и неравенство, уравнение и система, совокупность и система и т. д.) называются

равносильными, если множества их решений совпадают.

Нетрудно понять важность определения 19. По сути оно означает, что при решении уравнения (неравенства, системы, совокупности) можно уравнение (неравенство, систему, совокупность) заменить на любой равносильный ему объект. Значит, возможно, что в некоторых случаях уравнение сведётся к некоторому количеству более простых уравнений или неравенств, которые мы решать умеем.

Виден и основной недостаток этого определения: из него совершенно неясно, откуда брать эти самые уравнения и неравенства, равносильные данному. Или, что почти одно и то же, неясно, как установить равносильность внешне различных объектов (например, уравнения $\sin x = x$ и $2^x = \frac{1}{x+1}$ равносильны).

К счастью, в ряде случаев равносильность может быть установлена весьма просто. Эти случаи записываются в виде не очень сложных для понимания и доказательства теорем, многие из которых вам известны из курса средней школы. По сути любое «школьное» решение уравнения или неравенства представляет собой последовательное применение этих равносильностей; мы к ним настолько привыкли, что применяем их не задумываясь. В настоящей лекции посмотрим на них повнимательнее, они того заслуживают.

Теорема 5. Следующие условия равносильны:

- а) $f(x) > g(x)$ и $g(x) < f(x)$;
- б) $f(x) = g(x)$ и $g(x) = f(x)$;
- в) $f(x) \geq g(x)$ и $\begin{cases} f(x) > g(x) \\ f(x) = g(x) \end{cases}$;
- г) $f(x) = g(x)$ и $f(x) + h(x) = g(x) + h(x)$, если область определения функции h содержит пересечение областей определения функций f и g ;
- д) $f(x) < g(x)$ и $f(x) + h(x) < g(x) + h(x)$, если область определения функции h содержит пересечение областей определения функций f и g ;
- е) $f(x) = g(x)$ и $f(x) - g(x) = 0$;

ё) $f(x) > g(x)$ и $f(x) - g(x) > 0$;

ж) $f(x) = g(x)$ и $h(x) \cdot f(x) = h(x) \cdot g(x)$, если область определения функции h содержит пересечение областей определения функций f и g и $h(x) \neq 0$ при всех x из этого пересечения;

з) $f(x) > g(x)$ и $h(x) \cdot f(x) > h(x) \cdot g(x)$, если область определения функции h содержит пересечение областей определения функций f и g и $h(x) > 0$ при всех x из этого пересечения;

и) $f(x) > g(x)$ и $h(x) \cdot f(x) < h(x) \cdot g(x)$, если область определения функции h содержит пересечение областей определения функций f и g и $h(x) < 0$ при всех x из этого пересечения.

Замечание. Неравенство в пунктах (а), (д), (ё), (з), (и) теоремы 5 может быть нестрогим; из экономии места мы стали выписывать соответствующие равносильности. Также в пункте в) теоремы 5 можно поменять знак неравенства на противоположный.

Теорему 5 мы не будем доказывать; все её пункты очевидны. Однако обратим внимание, что ограничения, налагаемые на функцию h весьма существенны, игнорировать их нельзя. К сожалению, про это зачастую забывают и тогда вроде бы очевидные преобразования приводят к неверному ответу. Приведём два примера.

Пример 1. Уравнение $\sqrt{x+4} + \frac{1}{x} = 2 + \frac{1}{x}$ не равносильно уравнению $\sqrt{x+4} = 2$ (хотя казалось бы, пункт а) говорит об обратном), поскольку число 0 является решением второго уравнения, но не является решением первого. Причина ясна: функция h не определена в точке 0, а эта точка входит в область допустимых значений второго уравнения.

К этому примеру стоит добавить следующее.

Во-первых, преобразование указанного типа может всё же оказаться равносильным даже если функция h не такая, как того требует теорема: например, уравнения $\sqrt{x+3} + \frac{1}{x} = 2 + \frac{1}{x}$ и $\sqrt{x+3} = 2$ равносильны

(у обоих уравнений единственный корень $x = 1$). Однако в общем случае у нас нет гарантии, что равносильность будет соблюдена. Иногда здесь может помочь проверка, т. е. подстановка полученных корней в исходное уравнение (ниже мы поясним, в чём здесь дело). Однако ясно, что если корней много, или корни — иррациональные числа, с которыми трудно работать, то такая проверка затруднительна, а если корней бесконечно много (как часто бывает в неравенствах, а иногда и в уравнениях), она невозможна в принципе. Поэтому мы советуем по возможности осуществлять *только равносильные* преобразования. Например, в примере 1 равносильным был бы переход к системе
$$\begin{cases} \sqrt{x+4} = 2 \\ x \neq 0 \end{cases}.$$

Во-вторых, давайте осуществим похожее действие в чуть изменённом уравнении: $\sqrt{x-3} + \frac{1}{x} = 2 + \frac{1}{x} \Leftrightarrow \sqrt{x-3} = 2$. Оказывается, оно совершенно законно: изменилась область определения заключительного уравнения, и теперь она не включает те точки, где функция h не определена. Вывод: даже в случае простейших переходов следует приглядываться за их равносильностью.

Пример 2. Неравенство $\sqrt{x+4} \cdot (x+2) > 0$ не равносильно неравенству $\sqrt{x+4} > 0$ (хотя проведено преобразование, указанное в пункте (ж) теоремы 5), поскольку, например, число -3 является решением второго неравенства и не является решением первого: дело в том, что функция $h = \frac{1}{x+2}$ (именно на это выражение домножались обе части исходного неравенства) положительна не на всей области определения заключительного неравенства. Напротив, неравенство $\sqrt{x+4} \cdot (x+12) > 0$ равносильно неравенству $\sqrt{x+4} > 0$, прямо по пункту (ж): при $x \geq -4$ функция $h = \frac{1}{x+12}$ строго положительна.

Приведём ещё несколько примеров якобы равносильных переходов, на деле равносильными не являющимися.

Пример 3. «Взаимное уничтожение слагаемых». Неравенство

$x + \sqrt{x} - \sqrt{x} \leq 10$ не равносильно неравенству $x \leq 10$.

Пример 4. «Сокращение дробей». Неравенство $\frac{(x+1)^2}{x(x+1)} \geq 0$ не равносильно неравенству $\frac{x+1}{x} \geq 0$.

Пример 5. «Применение свойств корней». Уравнение $(\sqrt{x+1})^2 = x^2 - 5$ не равносильно уравнению $x + 1 = x^2 - 5$.

Пример 6. «Использование формул логарифмических преобразований». Неравенства $\log_2(x-2)(x-3) > 1$ и $\log_2(x-2) + \log_2(x-3) > 1$ не равносильны.

Пример 7. «Использование формул тригонометрии». Уравнения $\operatorname{tg}\left(x - \frac{\pi}{3}\right) = \frac{\sqrt{3}}{3}$ и $\frac{\operatorname{tg}x - \operatorname{tg}\frac{\pi}{3}}{1 + \operatorname{tg}x \cdot \operatorname{tg}\frac{\pi}{3}} = \frac{\sqrt{3}}{3}$ не равносильны.

Упражнение 29. Проанализируйте примеры 3 — 7 и определите причины неравносильности переходов в них; исправьте проведённые в примерах преобразования так, чтобы они стали равносильными.

Помимо равносильностей, указанных в теореме 5, вы со школы знаете более сложные равносильные переходы. Напомним их.

Теорема 6. Имеют место следующие равносильные переходы

$$\begin{aligned} \text{а) } f(x) \cdot g(x) = 0 &\Leftrightarrow \left[\begin{cases} f(x) = 0 \\ x \in D(g) \\ g(x) = 0 \\ x \in D(f) \end{cases} \right]; \\ \text{б) } f(x) \cdot g(x) > 0 &\Leftrightarrow \left[\begin{cases} f(x) > 0 \\ g(x) > 0 \\ f(x) < 0 \\ g(x) < 0 \end{cases} \right]; \end{aligned}$$

$$\begin{aligned} \text{в)} \quad f(x) \cdot g(x) < 0 &\Leftrightarrow \left[\begin{cases} f(x) > 0 \\ g(x) < 0 \\ f(x) < 0 \\ g(x) > 0 \end{cases} \right] ; \\ \text{г)} \quad \frac{f(x)}{g(x)} < 0 &\Leftrightarrow f(x) \cdot g(x) < 0; \quad \text{д)} \quad \frac{f(x)}{g(x)} > 0 \Leftrightarrow f(x) \cdot g(x) > 0. \end{aligned}$$

Замечание. В пунктах (б) и (в) можно все неравенства одновременно сделать нестрогими: равносильность сохранится. Однако в пунктах (г) и (д) полагать неравенства нестрогими нельзя; вывод: если мы желаем перейти от дроби к произведению в случае нестрогого неравенства следует отдельно рассматривать строгое неравенство и уравнение, т. е. предварительно воспользоваться равносильностью теоремы 5 (пункт в).

Доказательство теоремы 6 мы также опустим ввиду его очевидности. Отметим также тот факт, что пункт (а) теоремы 6 на школьном уровне формулируется так: «произведение равно нулю тогда и только тогда, когда хотя бы один из множителей равен нулю, а остальные множители при этом определены». Мы не случайно выделили последнюю часть фразы, не слишком грамотные люди её зачастую теряют, а как результат, получают посторонние корни. Пункты (б) и (в) в школе формулируются соответственно фразами «произведение двух множителей положительно тогда и только тогда, когда множители имеют одинаковые знаки» и «произведение двух множителей отрицательно тогда и только тогда, когда множители имеют разные знаки».

3. Равносильные переходы в типичных уравнениях и неравенствах

В этом разделе мы рассмотрим ряд равносильных преобразований, которые помогают решать простейшие уравнения и неравенства. В средней школе эти равносильности выделяются не всегда, разве что в специализированных математических классах. Нам же представляется разумным здесь их привести. Начнём с равносильностей в уравнениях и неравенствах, содержащих знак модуля.

Теорема 7. Имеют место следующие равносильные переходы

$$\begin{aligned} \text{а) } |f(x)| = g(x) &\Leftrightarrow \left[\begin{array}{l} \left\{ \begin{array}{l} f(x) = g(x) \\ f(x) \geq 0 \end{array} \right. \\ \left\{ \begin{array}{l} -f(x) = g(x) \\ f(x) < 0 \end{array} \right. \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} f(x) = \pm g(x) \\ g(x) \geq 0 \end{array} \right. ; \\ \text{б) } |f(x)| > g(x) &\Leftrightarrow \left[\begin{array}{l} f(x) > g(x) \\ f(x) < -g(x) \end{array} \right. ; \\ \text{в) } |f(x)| < g(x) &\Leftrightarrow -g(x) < f(x) < g(x) \Leftrightarrow \left\{ \begin{array}{l} f(x) < g(x) \\ f(x) > -g(x) \end{array} \right. . \end{aligned}$$

Доказательство. а) Первый переход изучается во всех школах без исключения — это просто определение модуля. Второй переход этого пункта доказывается разбором случаев, а именно: если число x таково, что $g(x) < 0$, то второе неравенство системы неверно, и, значит, такое число решением системы не является; не является оно и решением уравнения с модулем: модуль никакого числа не может быть отрицательным. Пусть теперь $g(x) \geq 0$. Тогда система сводится к своему первому уравнению: $f(x) = \pm g(x)$. Ясно, что такие x являются решениями уравнения с модулем: $|f(x)| = |\pm g(x)| = |g(x)| = g(x)$, а других решений уравнение иметь не может в силу того же определения модуля.

б) Второй переход очевиден для тех x , при которых $g(x) \geq 0$. Пусть x — такое число, что $g(x) < 0$. Тогда неравенство с модулем выполняется всегда, лишь бы в точке x была определена функция f . Это же касается и

совокупности из пункта (б): из двух условий $f(x)$ меньше фиксированного положительного числа и $f(x)$ больше некоторого отрицательного числа хотя бы одно всегда выполняется.

в) Доказывается аналогично пункту (б).

Упражнение 30. Сформулируйте и докажите аналогичные равносильные переходы для неравенств $|f(x)| \geq g(x)$ и $|f(x)| \leq g(x)$.

Замечание. В принципе при решении уравнений и неравенств, о которых говорится в теореме 7, вполне можно обойтись без приведённых в теореме равносильностей — достаточно просто определения модуля. Но применение этих равносильностей может существенно сократить работу. Например, решая уравнение $|x^3 + x^2 - 2x - 1| = x^3$ по определению модуля, мы будем вынуждены решать кубическое неравенство $x^3 + x^2 - 2x - 1 > 0$, а это непросто. Применяя же равносильность из теоремы 6 мы приходим к гораздо более простой задаче: найти неотрицательные корни уравнений $x^2 - 2x - 1 = 0$ и $2x^3 + x^2 - 2x - 1 = 0$.

Перейдём к иррациональным уравнениям и неравенствам. Чаще всего для их решения приходится возводить каждую из частей соответствующего уравнения (неравенства) возводить в натуральную степень — переход в ряде случаев не равносильный. Следующая теорема (доказательство которой мы не приводим) учит правильно осуществлять этот переход.

Теорема 8. Пусть $n \in \mathbf{N}$. Тогда имеют место следующие равносильные переходы

а) $f(x) = g(x) \Leftrightarrow f^{2n-1}(x) = g^{2n-1}(x)$, в частности $\sqrt[n]{f(x)} = g(x) \Leftrightarrow f(x) = g^{2n-1}(x)$;

б) $f(x) > g(x) \Leftrightarrow f^{2n-1}(x) > g^{2n-1}(x)$, в частности $\sqrt[n]{f(x)} > g(x) \Leftrightarrow f(x) > g^{2n-1}(x)$ и $f(x) > \sqrt[n]{g(x)} \Leftrightarrow f^{2n-1}(x) > g(x)$;

в) $f(x) = g(x) \Leftrightarrow f^{2n}(x) = g^{2n}(x)$, если обе функции $f(x)$ и $g(x)$ неотрицательны;

г) $f(x) > g(x) \Leftrightarrow f^{2n}(x) > g^{2n}(x)$, если обе функции $f(x)$ и $g(x)$

неотрицательны;

$$\begin{aligned} \text{д) } \sqrt[n]{f(x)} < g(x) &\Leftrightarrow \begin{cases} f(x) < g^{2n}(x) \\ f(x) \geq 0 \\ g(x) > 0 \end{cases} ; \\ \text{е) } \sqrt[n]{f(x)} > g(x) &\Leftrightarrow \begin{cases} \begin{cases} g(x) < 0 \\ f(x) \geq 0 \end{cases} \\ \begin{cases} g(x) \geq 0 \\ f(x) > g^{2n}(x) \end{cases} \end{cases} . \end{aligned}$$

Замечание. Равносильности пунктов (а) и (б) говорят лишь о том, что и уравнение, и неравенство можно возводить в любую нечётную степень (каждую часть отдельно) без ограничений. Возведение же в чётную степень допустимо только в том случае, если обе части уравнения (неравенства) неотрицательны — именно об этом пункты (в) и (г). Пункты (д) и (е) — следствие двух предыдущих; эти два пункта наиболее важны для случая иррациональных уравнений и неравенств, но специально их запоминать нет нужды: из пунктов (в) и (г) вытекает необходимость рассматривать случаи, когда обе части неотрицательны (и тогда можно заниматься возведением в степень) и когда это не так; последний случай анализируется тривиально.

Упражнение 31. Сформулируйте и докажите аналогичные равносильные переходы для неравенств $\sqrt[n]{f(x)} \geq g(x)$ и $\sqrt[n]{f(x)} \leq g(x)$.

Перейдём к показательным уравнениям и неравенствам. Они относятся к типу наиболее простых, изучаемых в школе. Происходит это потому, что показательная функция принимает только положительные значения, поэтому на неё можно смело умножать (или делить) обе части уравнения (неравенства), сокращать дроби, и т. д. Сложность могут вызвать только те случаи, где переменная имеется и в основании степени и в её показателе. Однако и здесь всё не так сложно. Просто нужно понимать, что запись $f(x)^{g(x)}$ трактуется как $e^{g(x) \ln f(x)}$ и только так (с точностью до основания, по которому берётся логарифм — оно может быть любым положительным, отличным от 1 числом), поэтому $f(x)$ обязана быть положительной (точ-

нее, те значения x , в которых $f(x) \leq 0$ не входят в область определения функции $f(x)^{g(x)}$. Кроме того, если $f(x) = 1$, то функция $f(x)^{g(x)}$ не может рассматриваться как показательная (напомним, что в определении показательной функции на основание накладывается ограничение не равняться 1). Конечно, последний случай анализируется легко, однако забывать его нельзя, а равносильности здесь таковы.

Теорема 9. Имеют место следующие равносильные переходы

$$\begin{aligned} \text{а) } (f(x))^{g(x)} = (f(x))^{h(x)} &\Leftrightarrow \left[\begin{aligned} &\begin{cases} g(x) = h(x) \\ f(x) > 0 \\ f(x) \neq 1 \end{cases} \\ &\begin{cases} x \in D(g) \cap D(h) \\ f(x) = 1 \end{cases} \end{aligned} \right] ; \\ \text{б) } (f(x))^{g(x)} > (f(x))^{h(x)} &\Leftrightarrow \left[\begin{aligned} &\begin{cases} g(x) > h(x) \\ f(x) > 1 \end{cases} \\ &\begin{cases} g(x) < h(x) \\ 0 < f(x) < 1 \end{cases} \end{aligned} \right] . \end{aligned}$$

Замечание 1. Для нестрого неравенства этого типа удачной равносильности нет, поэтому при его решении разумно сперва воспользоваться равносильностью из теоремы 5 (пункт в), а затем теоремой 9.

Замечание 2. Равносильности, приведённые в теореме 9 будут, конечно, верны и для случая, когда $f(x)$ — положительная константа. В этом случае их вид упрощается: в совокупности остаётся только одна система, а в этой системе только одно уравнение (неравенство). Вы получите равносильность, хорошо известную вам по школьному курсу.

Доказательство. а) Пусть число x таково, что имеет место равенство $(f(x))^{g(x)} = (f(x))^{h(x)}$. Тогда $f(x) > 0$ (иначе x не попадает в область допустимых значений уравнения). Если $f(x) \neq 1$, то функции в обеих частях уравнения — показательные, каждое своё значение принимают ровно один раз, поэтому $g(x) = h(x)$. Если же $f(x) = 1$, то в обеих частях уравнения

стоят 1, если только $x \in D(g) \cap D(h)$. Таким образом, выполнена совокупность пункта (а). Обратно, пусть для некоторого числа x выполнена указанная совокупность. Если верна первая система этой совокупности, то $(f(x))^{g(x)} = (f(x))^{h(x)}$ по свойству показательной функции. Если выполнена вторая её система, то $(f(x))^{g(x)} = (f(x))^{h(x)}$ превращается в равенство $1 = 1$, т. е. тоже становится очевидным.

б) Пусть число x зафиксировано. Возможно 4 случая: 1) $f(x) \leq 0$ или $f(x)$ не определено — тогда не определено выражение $f(x)^{g(x)}$, и совокупность тоже не выполнена; 2) $f(x) = 1$ — тогда ложна и левая и правая часть равносильности (б); 3) $f(x) > 1$ — тогда неравенство равносильно первой системе совокупности; 4) $0 < f(x) < 1$ — неравенство равносильно второй системе. Теорема доказана.

Теперь рассмотрим некоторые равносильности в уравнениях и неравенствах, содержащими логарифмическую функцию. Рассмотрим только случай, когда в обеих частях уравнения (неравенства) стоят логарифмы по одному и тому же (возможно, переменному) основанию. Тогда справедлива теорема.

Теорема 10. Имеют место следующие равносильные переходы

$$\begin{aligned} \text{а) } \log_{f(x)} g(x) = \log_{f(x)} h(x) &\Leftrightarrow \begin{cases} g(x) = h(x) \\ h(x) > 0 \\ f(x) > 0 \\ f(x) \neq 1 \end{cases} \Leftrightarrow \begin{cases} g(x) = h(x) \\ g(x) > 0 \\ f(x) > 0 \\ f(x) \neq 1 \end{cases} ; \\ \text{б) } \log_{f(x)} g(x) > \log_{f(x)} h(x) &\Leftrightarrow \begin{cases} \begin{cases} g(x) > h(x) > 0 \\ f(x) > 1 \end{cases} \\ 0 < g(x) < h(x) \\ 0 < f(x) < 1 \end{cases} . \end{aligned}$$

Доказательство теоремы 10 проводится также, как и доказательство теоремы 9 (разбором нескольких случаев), поэтому мы его опускаем. Обратим внимание на то, что мы в теореме ничего не говорим про область допустимых значений. На самом деле, она учтена, только неявно. Внимательный студент и даже школьник легко это проверит.

Замечание. В пункте (а) приведены две совокупности, каждая из которых равносильна логарифмическому уравнению. Они абсолютно идентичны, можно было обойтись любой одной. Не сделано это для того, чтобы явно указать на тот факт, что мы имеем возможность выбора, какое из двух неравенств $g(x) > 0$ или $h(x) > 0$ не решать (конечно, которое потруднее).

Упражнение 32. Что изменится в пункте (б) теоремы 10, если исходное логарифмическое неравенство будет нестрогим?

Из многочисленных тригонометрических равносильностей приведём только пару наиболее простых. Но при этом доказательство их оставим для самостоятельной работы.

Теорема 11.

$$\begin{aligned} \text{а) } \sin(f(x)) = \sin(g(x)) &\Leftrightarrow \begin{cases} f(x) = g(x) + 2\pi k, & k \in \mathbf{Z} \\ f(x) = \pi - g(x) + 2\pi k, & k \in \mathbf{Z} \end{cases}; \\ \text{б) а) } \cos(f(x)) = \cos(g(x)) &\Leftrightarrow \begin{cases} f(x) = g(x) + 2\pi k, & k \in \mathbf{Z} \\ f(x) = -g(x) + 2\pi k, & k \in \mathbf{Z} \end{cases}. \end{aligned}$$

Упражнение 33. Докажите теорему 11.

Упражнение 34. Выведите аналогичные равносильности для равенства тангенсов двух функций и котангенсов двух функций.

4. Переходы к следствию. Равносильность и область допустимых значений

К сожалению, переход от уравнения (неравенства) к равносильному зачастую оказывается сложным. Поэтому в ряде случаев от равносильности отказываются. Разумеется, делать это следует очень осторожно, чтобы во-первых, не потерять корней, а во-вторых, чтобы не получить посторонних.

Для решения первой из указанных задач очень полезно понятие следствия.

Определение 20. Уравнение (неравенство, система, совокупность) А называется *следствием* уравнения (неравенства, системы, совокупности) Б, если множество решений А содержит множество решений Б.

Пример 8. Уравнение $f(x) = g^2(x)$ является следствием уравнения $\sqrt{f(x)} = g(x)$, поскольку все корни второго уравнения являются также и корнями первого (в действительности корнями первого уравнения будут в точности все корни совокупности уравнений $\sqrt{f(x)} = g(x)$ и $\sqrt{f(x)} = -g(x)$).

Пример 9. Уравнение $f(x) = g(x)$ является следствием уравнения $f(x) + h(x) = g(x) + h(x)$, но не наоборот (убедитесь в этом самостоятельно).

Пример 10. Совокупность $\begin{cases} f(x) = 0 \\ g(x) = 0 \end{cases}$ является следствием уравнения $f(x) \cdot g(x) = 0$ (в школе это звучало так: «если произведение двух множителей равно нулю, то хотя бы один из них равен нулю»).

Пример 11. Неравенство $f(x) \leq g(x)$ является следствием двойного неравенства $f(x) \leq h(x) \leq g(x)$ — это совершенно очевидно.

Заметим, что непосредственно из определения 20 следует очевидный факт:

Утверждение. Два уравнения (два неравенства, уравнение и неравенство, уравнение и система, совокупность и система и т. д.) являются равносильными, если и только если каждое из них является следствием другого.

Утверждение «Б является следствием А» может также выражаться фразами «из А следует Б», «условие А достаточно для Б», «условие Б необходимо для А», «если А, то Б» и т. д. Просьба обратить особое внимание на слова «необходимо» и «достаточно»: их смысл часто путают.

Как видно из определения, если при решении уравнения (неравенства, системы, совокупности) мы в какой-то момент перешли к следствию, то потеря корня нам не грозит, а вот получение посторонних корней вполне и вполне вероятно. Это, конечно, не очень хорошо, но не так опасно, как обратная ситуация, когда посторонних корней не будет точно, но при этом возможна потеря реальных. Собственно говоря, эта обратная ситуация просто противопоказана в математике; следует твёрдо запомнить (если это ещё не выучено со школы), что *переход не к следствию не допускается никогда*. Переход же к следствию (не являющийся равносильным) допустим, но следует помнить, что после него ещё нужно провести отбор корней (например, путём проверки).

При записи решения, как правило, придерживаются следующей договорённости: если уравнения (неравенства, системы) записаны друг под другом, то это то же самое, как если бы между ними стоял знак \Leftrightarrow , т. е. имеет место равносильный переход. Переход к следствию всегда оговаривается особо (иногда просто ставится знак \Rightarrow). Наличие хотя бы одного такого перехода требует в конце решения дополнительного анализа на предмет выявления посторонних корней.

В качестве примера правильной работы с переходом к следствию рассмотрим решение следующей задачи:

Пример 12. Решить систему:
$$\begin{cases} x^4 + x^3 - 3x^2 \leq 5^{x+1} - 4^{-x} \\ 5^{x+1} - 4^{-x} \leq x^3 - x^2 - 1 \end{cases}.$$

Решение. Система имеет вид
$$\begin{cases} f(x) \leq g(x) \\ g(x) \leq h(x) \end{cases}.$$
 Перейдём к следствию $f(x) \leq h(x)$, т. е. к неравенству $x^4 + x^3 - 3x^2 \leq x^3 - x^2 - 1$. Решим полученное неравенство-следствие:

$$x^4 + x^3 - 3x^2 \leq x^3 - x^2 - 1$$

$$x^4 - 2x^2 + 1 \leq 0$$

$$(x^2 - 1)^2 \leq 0$$

$$x^2 - 1 = 0$$

$$x = \pm 1.$$

Мы получили два возможных корня данной системы. Поскольку в решении имел место переход к следствию, не исключена возможность того, что среди корней будут посторонние. Осуществим проверку и увидим, что:

1) число $x = -1$ является корнем: $f(x) = g(x) = h(x) = -3$;

2) число $x = 1$ корнем не является: неравенство $5^{1+1} - 4^{-1} \leq 1^3 - 1^2 - 1$ неверно.

Ответ: $x = -1$.

В заключении лекции скажем несколько слов об области допустимых значений. Ряд учебников рекомендует решение любого уравнения или неравенства начинать с поиска этого множества. Нам же представляется, что это не вполне хорошо, поскольку зачастую требует лишней и технически сложной работы. Да и панацеей от приобретения посторонних корней такая работа не является. Однако, в тех случаях, когда ОДЗ легко находится, пренебрегать им не следует. Причина тому такова: при сужении области значений переменной появляются новые равносильности, которые в общем случае места не имеют. Точнее, возникает новое понятие: равносильность на некотором множестве.

Определение 21. Пусть $A \subset \mathbf{R}$. Два уравнения I_1 и I_2 (два неравенства, уравнение и неравенство, уравнение и система, совокупность и система и т. д.) называются *равносильными на множестве A* , если всякое решение уравнения I_1 , принадлежащее множеству A , является также и решением уравнения I_2 и наоборот; иными словами, уравнения равносильны на множестве A , если совпадают множества их решений, пересечённые с множеством A .

Пример 13. Уравнение $\cos(\pi x) = \frac{1}{2}x + \frac{1}{3}$ и $x = \frac{1}{3}$ равносильны на множестве $[0; 1]$, но не равносильны на всей числовой прямой. В самом деле, на отрезке $[0; 1]$ функция $y = \cos(\pi x)$ монотонно убывает, а функция $y = \frac{1}{2}x + \frac{1}{3}$ монотонно возрастает, поэтому на этом отрезке уравнение не может иметь более одного корня, а число $x = \frac{1}{3}$ — очевидный корень. На всей же числовой прямой уравнение имеет несколько корней; это легко можно увидеть из графиков функций, стоящих в обеих частях уравнения, или обосновать таким образом: функция $h(x) = \cos(\pi x) - \left(\frac{1}{2}x + \frac{1}{3}\right)$ непрерывна, положительна в точке 0 и отрицательна, в точке -1 . Поэтому на отрезке $[-1; 0]$ она имеет хотя бы один корень, который является корнем уравнения $\cos(\pi x) = \frac{1}{2}x + \frac{1}{3}$, отличным от числа $\frac{1}{3}$.

Упражнение 35. Пусть $f(x)$ — функция, определённая на всей числовой прямой. Обязательно ли являются равносильными при $x \geq a$:

- а) уравнения $\sqrt{x-a} = f(x)$ и $x-a = f^2(x)$;
- б) неравенства $\sqrt{x-a} < f(x)$ и $x-a < f^2(x)$;
- в) неравенства $\sqrt{x-a} > f(x)$ и $x-a > f^2(x)$?

Ясно, что решая уравнение (неравенство, систему и т. д.), можно рассматривать только те значения переменной, которые лежат в ОДЗ этого уравнения (неравенства, системы). И зачастую на этом более узком множестве имеют место равносильности, в общем случае места не имеющие. Приведём два примера.

Пример 14. Решая неравенство $\log_2 x^2 \geq \log_2(4x-3)$ просто опустить логарифмы и перейти к неравенству $x^2 \geq 4x-3$, разумеется, нельзя: число -1 , например, удовлетворяет второму неравенству, но не удовлетворяет первому. (В действительности этот переход есть переход к следствию — равносильные переходы были разобраны в этой лекции ранее — а проверка невозможна, так как полученное квадратное неравенство имеет бесконечное множество решений). Однако, если мы заранее нашли ОДЗ, т. е. решили

систему $\begin{cases} x^2 > 0 \\ 4x - 3 > 0 \end{cases}$, переход был бы вполне допустим. Только нужно не забыть, что решив полученное неравенство, следует ответ (в нашем случае множество $(-\infty; 1) \cup (3; +\infty)$) пересечь с ОДЗ (т. е. с множеством $(\frac{3}{4}; +\infty)$) и получить верный ответ $x \in (3; +\infty)$.

Пример 15. Решить систему $\begin{cases} |5 - x| + |x^2 - 5| = 2 \\ \sqrt{7x - 12 - x^2} \geq 2^{x-3} \end{cases}$.

Очень не хочется решать неравенство системы: оно не относится к типичным «школьным» неравенствам, алгоритм решения сразу не виден. Решение же уравнения системы подразумевает раскрытие модулей, т. е. решение нескольких систем. Впечатление, что решение системы требует длинных выкладок. Но достаточно заметить, что ОДЗ системы (а у нас единственное ограничение: под знаком корня должно стоять неотрицательное число) — это отрезок $[3, 4]$, как задача становится почти устной. В самом деле, при $x \in [3, 4]$ оба подмодульных выражения первого уравнения положительны, поэтому уравнение равносильно (на ОДЗ) равенству $x^2 - x = 2$. Но оба корня последнего уравнения ОДЗ не принадлежат, поэтому уравнение на ОДЗ корней не имеет, и, значит, вся система решений также не имеет.

Вывод прост: если ОДЗ найдено, то уравнение (система, совокупность) может решиться проще, так как на меньшем множестве возникает больше равносильных переходов. Однако ещё раз обратим внимание на то, что равносильность порой оказывается кажущейся, а на самом деле имеет место переход к следствию (в лучшем случае). Сравните два примера.

Пример 16. Решить уравнение $\log_3 x + \log_3 (2x - 15) = 3$.

На области допустимых значений (легко видеть, что это бесконечный интервал $(7, 5; +\infty)$) имеет место равносильное преобразование: сумму логарифмов можно заменить логарифмом произведения. Получим уравнение $\log_3 (2x^2 - 15x) = 3$, откуда $2x^2 - 15x = 27$. Это уравнение имеет два корня $x_1 = 9$ и $x_2 = -1, 5$, причём первый входит в ОДЗ, а второй не входит,

т. е. является посторонним. Ответ: $x = 9$.

Как видно, в этом примере допустимо использование формулы суммы логарифмов.

Пример 17. Решить уравнение $\log_3 x(x - 8) = \log_3 (2x + 2)(x - 8)$.

ОДЗ данного уравнения — это множество решений системы
$$\begin{cases} x(x - 8) > 0 \\ (2x + 2)(x - 8) > 0 \end{cases}$$
. Решив систему, получаем, что ОДЗ — объединение двух бесконечных интервалов: $(-\infty; -1) \cup (8; +\infty)$. Попробуем применить ту же формулу суммы логарифмов, что и в предыдущем случае: $\log_3 x + \log_3 (x - 8) = \log_3 (2x + 2) + \log_3 (x - 8)$. И выясняется, что мы уже потеряли корень. В самом деле, число $x = -2$ является решением исходного уравнения, и не является решением последнего уравнения. А ведь точка $x = -2$ в ОДЗ входит!

Вывод: применение формулы суммы логарифмов в данном случае недопустимо; наиболее же простым способом решения является переход к уравнению-следствию $x(x - 8) = (2x + 2)(x - 8)$ и последующая проверка.

ЛЕКЦИЯ VI. ОСНОВЫ КОМБИНАТОРИКИ.

ОСНОВНЫЕ ПРИНЦИПЫ

1. Конечные и бесконечные множества

Комбинаторика — это раздел математики, изучающий конечные множества, т. е. такие, которые можно задать полным перечислением всех элементов. Разумеется, когда мы говорим, что можем перечислить все элементы некоторого множества, то предполагаем чисто теоретическую возможность сделать это. Например, перечислить все молекулы кислорода в этой комнате реально невозможно, однако любой физик скажет вам (вы это и так знаете со школы), что примерное количество молекул может быть подсчитано — такие формулы физикам известны. Это означает, что (конечно, чисто теоретически), молекулы можно пересчитать, присвоить каждой свой

порядковый номер. И, значит, множество молекул кислорода в комнате конечно. Другой пример, чисто математический. Рассмотрим множество корней уравнения $x^{13} + 7x^4 - 18x^3 + 2x = 1$. Вряд ли вы сможете указать хотя бы один элемент рассматриваемого множества, не говоря о том, чтобы найти все его корни. Даже вопрос о количестве корней уравнения сложен. Однако хорошо известно, что уравнение степени n не может иметь корней больше, чем n , значит, рассмотренное множество имеет не более 13 элементов и может быть (опять-таки теоретически) задано их перечислением. Таким образом, множество корней данного уравнения конечно. Считается конечным и пустое множество, хотя его элементы перечислить нельзя ввиду их отсутствия. Строгое же определение конечного множества таково.

Определение 22. Множество называется *конечным*, если оно или пусто или допускает биективное отображение на множество $\{1, 2, \dots, k\}$ при некотором $k \in \mathbb{N}$.

Иными словами, всякое конечное непустое множество может быть записано в виде $\{a_1, a_2, \dots, a_k\}$, где все a_i — различные элементы множества. Собственно, в такой форме записи явно просматривается биекция, о которой говорится в определении конечного множества, а именно, эта биекция ставит в соответствие элементу a_i число i ($1 \leq i \leq k$).

Легко видеть, что всякое множество может допускать биективное отображение на множество $\{1, 2, \dots, k\}$ не более, чем при одном значении k . Действительно, если бы для некоторого множества A нашлись бы биекции $h_1 : A \rightarrow \{1, 2, \dots, k\}$ и $h_2 : A \rightarrow \{1, 2, \dots, m\}$, где k и m — различные натуральные числа, то (см. упражнение 20) отображение $h_2 \circ h_1^{-1}$ было бы биекцией множества $\{1, 2, \dots, k\}$ на множество $\{1, 2, \dots, m\}$, что, очевидно, невозможно. На самом деле эту очевидность не так легко обосновать; в настоящем цикле лекций мы это обоснование опустим.

Таким образом, каждому непустому конечному множеству A ставится в соответствие некоторое натуральное число k — то самое для которого существует рассмотренная биекция. Смысл числа k — количество элементов

множества A . На языке математики это число называется *мощностью* множества A и обозначается $|A|$. Кроме того, полагают $|\{\emptyset\}| = 0$, что вполне естественно.

Множества, не являющиеся конечными, называются *бесконечными*. Здесь имеет место некоторое расхождение понятий бесконечности в математике и в обычной жизни. Например, говоря, что количество всевозможных разных шахматных партий бесконечно велико, мы подразумеваем, что ни человеку, ни компьютеру не под силу за разумное время перебрать все возможные последовательности ходов, образующих шахматную партию. Однако существуют шахматные правила, ограничивающие длину партии (в первую очередь правило 50 ходов), и так как в каждой позиции число ходов конечно, то с математической точки зрения различных шахматных партий тоже конечное число, хотя это число очень велико. Как уже говорилось в начале лекции, бесконечные множества мы рассматривать сейчас не будем.

Главный вопрос комбинаторики — вопрос о мощности конечного множества. Грубо говоря, если в задаче требуется установить количество элементов некоторого конечного множества — то эта задача по комбинаторике. Для решения этой задачи в комбинаторике введены специальные понятия, такие как сочетания, размещения, перестановки (о них пойдёт речь ниже) и выведены специальные формулы. Однако на самом деле без всех этих понятий и формул вполне можно обойтись; комбинаторные задачи успешно решаются просто исходя из здравого смысла, а применение формул позволяет просто сократить обоснование решения и запись ответа.

Кстати, об ответе. Ответом на вопрос «сколько» является, естественно, число. В случае комбинаторных задач — число натуральное или ноль (вероятно, именно рассмотрение задач комбинаторики в качестве основных задач математики привело к тому, что в ряде стран мира, например, во Франции, число ноль считается натуральным). Однако вычислять это число, находить его десятичную запись, в комбинаторных задачах не требуется. Если ответ записан в виде, который после выполнения конечного числа арифметических операций над числами оказывается натуральным числом в десятичной записи, то этого вполне достаточно, задача по комби-

наторике будет признана решённой. Например, допустима запись ответа в виде $\frac{13 \cdot 14 \cdot 15}{2 \cdot 3 \cdot 4}$ или $\sum_{i=1}^3 (i^2 - i)$. Также нет нужды расписывать обозначения количества сочетаний, перестановок и размещений, например достаточно записать просто $C_{24}^3 \cdot \overline{A}_{16}^5$ (что обозначают эти буквы и индексы при них будет пояснено ниже). Для современных студентов, «воспитанных на заданиях ГИА и ЕГЭ», требующих записи ответа в строго фиксированном виде тестовых технологиях, это может показаться странным. Однако по-иному в комбинаторике нельзя. Приведём пример известнейшей комбинаторной задачи.

Как гласит одна из легенд, древнеиндийскому мудрецу — изобретателю шахмат — была обещана следующая награда: за первое поле шахматной доски — одно зерно, за второе — 2 зерна, за третье — 4 и т. д., т. е. за каждое следующее поле полагалось в два раза больше зёрен, чем за предыдущее. Сколько зёрен должен был получить мудрец?

Решение задачи несложно. Так как всего полей на доске 64, то мудрецу причитается $1 + 2 + 4 + \dots + 2^{63}$ зёрен. По формуле суммы геометрической прогрессии это число равно $2^{64} - 1$, что и является ответом. Однако выписать десятичную запись полученного числа сможет не каждый: очень большой счёт. Да и калькулятор не поможет: разрядной сетки не хватит. Понятно, что этот счёт относится к арифметике, а к комбинаторике никакого отношения не имеет, и требовать его осуществления неразумно.

2. Основные принципы комбинаторики

Решение комбинаторных задач опирается на ряд простейших, почти очевидных фактов, которые называются основными принципами. Эти факты вам должны быть хорошо известны, однако вполне вероятно, что вы не заостряли на них внимание. Укажем их.

Принцип первый (принцип соответствия). Если существует биекция одного конечного множества на другое (т. е. если каждому элементу одного множества поставлен в соответствие ровно один элемент другого, причём разным элементам соответствуют разные), то эти множества содержат

равное число элементов и наоборот, если множества содержат поровну элементов, то между ними существует биекция.

Строгое доказательство этого принципа приводит к необходимости доказывать отсутствие биекции между множествами $\{1, 2, \dots, m\}$ и $\{1, 2, \dots, n\}$ при $m \neq n$. Технически это не очень просто, а сам факт достаточно очевидный, поэтому доказывать ни этот факт, ни сам принцип соответствия, мы не будем. Просто приведём пример.

Пример 18. В группе обучаются 25 студентов, из которых один умеет играть на гитаре; назовём его гитаристом. Будем называть *компанией* любое подмножество студентов группы, в котором более одного человека и будем называть компанию *музыкальной*, если в неё входит гитарист. Каких компаний больше и на сколько: музыкальных, или не музыкальных?

Решение. Рассмотрим некоторую не музыкальную компанию. Если в неё добавить гитариста, то компания станет музыкальной. Ясно, что мы получили отображение из множества всех не музыкальных компаний в множество музыкальных. Это отображение инъективно (поскольку разным не музыкальным компаниям соответствуют разные музыкальные), однако не сюръективно. В самом деле, добавляя гитариста в любую не музыкальную компанию, мы получим компанию из трёх человек по крайней мере, поэтому в музыкальную компанию из двух человек (гитарист плюс ещё кто-то) ни одна компания не отображается. Однако если построенное отображение рассматривать как отображение из множества всех не музыкальных компаний в множество всех музыкальных компаний, имеющих более двух членов, получится биекция (докажите), поэтому в силу принципа соответствия таких компаний поровну. Значит, музыкальных компаний больше на число всех компаний вида гитарист плюс ещё один участник. Этот ещё один участник может быть любым из 24 оставшихся студентов, и значит, таких компаний 24.

Ответ: музыкальных компаний больше на 24.

Упражнение 36. На окружности отметили 20 точек. Рассматриваются всевозможные многоугольники, все вершины которых являются отмеченными точками. Каких многоугольников среди них больше, 8-угольников или 12-угольников?

Принцип второй (теорема о сложении). Если конечные множества A_1, A_2, \dots, A_n попарно не пересекаются, то количество элементов в их объединении равняется сумме количеств элементов в каждом из множеств. (Математически это записывается формулой $\left| \bigcup_{i=1}^n A_i \right| = \sum_{i=1}^n |A_i|$.)

Доказательство этого принципа также вполне очевидно. Приведём его для случая $n = 2$ (для остальных n доказательство не меняется, просто усложняется запись функции φ).

Доказательство. Пусть $|A_1| = p$, $|A_2| = t$, где p и t — целые неотрицательные числа. Тогда $A_1 = \{a_1, a_2, \dots, a_p\}$ и $A_2 = \{b_1, b_2, \dots, b_t\}$. Так как $A_1 \cap A_2 = \emptyset$, то все элементы множества $A_1 \cup A_2 = \{a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_t\}$ различны. Рассмотрим функцию $\varphi : \{1, 2, \dots, p+t\} \rightarrow A_1 \cup A_2$ такую, что $\varphi(n) = a_n$ при $n \leq p$ и $\varphi(n) = b_{n-p}$ при $p < n \leq p+t$. Она является биекцией, поэтому $|A_1 \cup A_2| = p+t = |A_1| + |A_2|$. Теорема доказана.

У теоремы о сложении имеется и другая формулировка, эквивалентная той, которая была приведена. Связано это с тем, что практически любую задачу по комбинаторике можно свести к решению задачи о количестве способов осуществить некоторый выбор. Например, вопрос о том, сколько существует двухэлементных подмножеств заданного множества, эквивалентен вопросу сколькими способами можно выбрать два различных элемента из заданного множества. Поэтому многие учебники по комбинаторике написаны на языке числа способов выбора. Это довольно удобный и наглядный язык, мы тоже будем его время от времени использовать. Теорема о сложении на этом языке звучит так.

Теорема 12 (о сложении). Пусть имеется n различных объектов ($n \in \mathbf{N}$), из которых надо выбрать один. Пусть при этом i -ый объект может быть выбран k_i числом способов ($i = \overline{0, n}$). Тогда количество способов осуществить требуемый выбор равно $k_1 + k_2 + \dots + k_n$.

Ввиду простоты теоремы мы не будем приводить примера, её иллюстрирующего. Однако заметим, что для того, чтобы теорема «работала» необходимо, чтобы объединяемые множества не пересекались. Несколько позже мы покажем, как работать с пересекающимися множествами.

Принцип третий (теорема об умножении.) Пусть даны конечные множества A_1, A_2, \dots, A_n . Тогда количество элементов их декартова произведения равно произведению количеств элементов перемножаемых множеств. Математическая запись такова: $|A_1 \times A_2 \times \dots \times A_n| = \prod_{i=1}^n |A_i|$.

Доказательство. Докажем теорему только для случая $n = 2$. Полное доказательство проще всего провести методом математической индукции; о нём пойдёт речь позже. После изучения индукции доказательство теоремы об умножении для произвольного количества сомножителей становится просто упражнением.

Итак, пусть имеется два множества $A = \{a_1, \dots, a_k\}$ и $B = \{b_1, \dots, b_l\}$ и докажем, что $|A \times B| = kl$. Заметим, что $A \times B = \bigcup_{i=1}^k (\{a_i\} \times B)$, и что множества $\{a_i\} \times B$ попарно не пересекаются: при разных i у пар — элементов множества $\{a_i\} \times B$ различны первые элементы. Тогда по теореме о сложении $|A \times B| = \sum_{i=1}^k |\{a_i\} \times B|$. Но всякое множество $\{a_i\} \times B$ содержит столько же элементов, сколько и множество B , т. е. l штук. Тогда $|A \times B| = \sum_{i=1}^k l = kl$, что и требуется доказать.

Как и в случае с теоремой сложения приведём формулировку этого принципа на языке выбора.

Теорема 13 (об умножении). Пусть имеется n объектов ($n \in \mathbf{N}$), и необходимо выбрать их все (каждый объект в количестве 1 штука). Пусть при этом i -ый объект может быть выбран k_i числом способов ($i = \overline{0, n}$). Тогда количество способов осуществить выбор равно $k_1 \cdot k_2 \cdot \dots \cdot k_n$.

Пример 19. В студенческой столовой предлагается 3 первых блюда, 4 вторых и 8 третьих. Сколькими способами студент может выбрать себе полный обед (состоящий из первого, второго и третьего блюда)?

Решение. Всякий полный обед есть элемент декартова произведения трёх множеств: первых, вторых и третьих блюд. По теореме об умножении количество полных обедов равно $3 \cdot 4 \cdot 8 = 96$. Это и есть число способов выбрать полный обед.

Ответ: 96-ю способами.

Заметим, что в отличие от теоремы о сложении, множества A_i в теореме об умножении могут иметь непустое пересечение и даже совпадать; например, если требуется подсчитать количество трёхзначных чисел, все цифры которых нечётны, то мы можем воспользоваться теоремой об умножении для множеств $A_1 = A_2 = A_3 = \{1, 3, 5, 7, 9\}$ (A_i — множество цифр, которые могут стоять в i -ом разряде числа). Разумеется, ответом будет являться число $5 \cdot 5 \cdot 5 = 125$.

Типичной ошибкой слабых студентов (и школьников) является неверное понимание, в каких случаях следует применить теорему о сложении, а в каких теорему об умножении. Существует мнемоническое правило, которое позволяет разобраться в этом вопросе. Связано это со следующим наблюдением.

Когда мы выбираем один элемент из нескольких, мы употребляем союз «или» (элемент может быть «или» первым, «или» вторым, «или» третьим и т. д.). Когда же мы выбираем все элементы по разу, то употребляем союз «и» (выберем «и» первый элемент, «и» второй, «и» третий...). Таким обра-

зом, если в тексте задачи стоит союз «или», то надо применять теорему о сложении, если союз «и» — теорему об умножении. Это и есть мнемоническое правило, о котором идёт речь.

Для иллюстрации сказанного приведём пример, описанный в художественном произведении. Два не очень грамотных человека обсуждают десятискоростной велосипед. И приходят к выводу, что на самом деле он имеет не 10, а только 7 скоростей. Их рассуждения таковы: «Скорость велосипеда определяется положением велосипедной цепи. Она держится на двух втулках, передней и задней. На передней втулке у цепи 5 разных положений, на задней — только два. Так как $2 + 5 = 7$, цепь может принимать 7 разных положений, соответственно, и скоростей у велосипеда тоже 7». Ошибка в их рассуждении совершенно очевидна, и связана как раз с тем, что положение цепи определяется её положением на передней «и» на задней втулке, поэтому применять следует теорему об умножении, а не теорему о сложении.

Упражнение 37. В 11-а классе средней школы обучается 13 юношей и 18 девушек, а в 11-б классе — 16 юношей и 12 девушек. Для приветственного слова директор желает выбрать по одному ученику из этих двух классов, причём разного пола. Сколькими способами он может сделать такой выбор?

Принцип четвёртый (формула включений — исключений.) Пусть имеется n конечных множеств, возможно, пересекающихся A_1, A_2, \dots, A_n . Тогда количество элементов в объединении всех этих множеств может быть подсчитано следующим образом: надо сложить количества элементов всех множеств, из суммы вычесть все количества элементов их попарных пересечений, затем добавить количества элементов всех пересечений из трёх множеств, вычесть — из четырёх и т. д. вплоть до учёта пересечения всех n множеств (которое надо будет вычесть при чётном n и прибавить при нечётном).

Математическая запись формулы в общем виде очень громоздка и не наглядна. Мы приведём эту запись (равно как и доказательство формулы) только для двух частных случаев — $n = 2$ и $n = 3$. Как правило, в задачах только эти случаи и имеют место. Причина тому проста: при увеличении n резко возрастает число слагаемых (на самом деле оно равно $2^n - 1$) и подсчёт становится нереально тяжёлым. Заметим также, что при $n = 2$ и $n = 3$ теорема имеет наглядную картинку на диаграммах Венна (иногда их называют *кругами Эйлера*), при больших n наглядность теряется. Итак, вот математическая формулировка частных случаев формулы включений — исключений.

Теорема 14 (*формула включений — исключений*). Пусть A , B и C — конечные множества. Тогда

$$\text{I.} \quad |A \cup B| = |A| + |B| - |A \cap B|.$$

$$\text{II.} \quad |A \cup B \cup C| = |A| + |B| + |C| - (|A \cap B| + |A \cap C| + |B \cap C|) + |A \cap B \cap C|.$$

Доказательство I. Заметим, что $A \cup B = (A \setminus B) \cup B$ и что множества $A \setminus B$ и B имеют пустое пересечение. Тогда по теореме о сложении $|A \cup B| = |A \setminus B| + |B|$. Далее, $A = (A \setminus B) \cup (A \cap B)$ и множества $A \setminus B$ и $A \cap B$ также имеют пустое пересечение. По теореме о сложении имеем $|A| = |A \setminus B| + |A \cap B|$, откуда $|A \setminus B| = |A| - |A \cap B|$. Окончательно имеем

$$|A \cup B| = |A \setminus B| + |B| = |A| - |A \cap B| + |B|,$$

что и требуется доказать.

II. Обозначим множество $A \cup B$ через D . Тогда $A \cup B \cup C = D \cup C$, согласно пункту I этой теоремы $|D \cup C| = |D| + |C| - |D \cap C|$ и по той же причине $|D| = |A| + |B| - |A \cap B|$. Кроме того, $D \cap C = (A \cup B) \cap C = (A \cap C) \cup (B \cap C)$, поэтому всё по тому же пункту I

$$\begin{aligned} |D \cap C| &= |A \cap C| + |B \cap C| - |(A \cap C) \cap (B \cap C)| = \\ &= |A \cap C| + |B \cap C| - |A \cap B \cap C|. \end{aligned}$$

Окончательно имеем

$$\begin{aligned} |A \cup B \cup C| &= |D \cup C| = |D| + |C| - |D \cap C| = \\ &= |A| + |B| - |A \cap B| + |C| - (|A \cap C| + |B \cap C| - |A \cap B \cap C|), \end{aligned}$$

откуда после раскрытия скобок и следует утверждение пункта II.

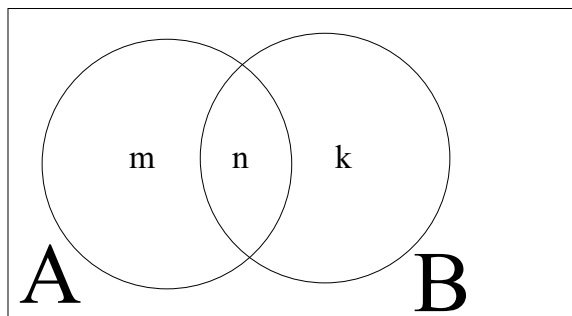


Рис. 1. К доказательству теоремы 14 (I)

Разумеется, доказательство этих пунктов может быть проведено и с помощью диаграмм Венна. Например, первый пункт доказывается так: обозначим через m , n и k количества элементов множеств $A \setminus B$, $A \cap B$ и $B \setminus A$ соответственно (см. рис. 1). Тогда $|A| = m + n$, $|B| = k + n$, $|A \cup B| = m + n + k$

и требуемое равенство доказывается элементарным подсчётом: проверяется справедливость формулы $m + n + k = (m + n) + (k + n) - n$. Общий же случай теоремы (при $n > 3$) доказывается методом математической индукции.

В качестве примера на применение формулы включений-исключений приведём такую задачу.

Пример 20. Сколько существует трёхзначных чисел, не делящихся ни на 3, ни на 7, ни на 11?

Решение. Общее количество трёхзначных чисел 900 (наибольшее трёхзначное число — это 999, и из этого числа нужно вычесть 99 первых натуральных чисел, не являющихся трёхзначными. Пусть A , B и C — множества всех трёхзначных чисел, делящихся соответственно на 3, 7 и 11. Тогда нас интересует число $900 - |A \cup B \cup C|$. Так как $A = \{102 = 3 \cdot 34, 3 \cdot 35, \dots, 999 = 3 \cdot 333\}$, то $|A| = 333 - 33 = 300$. Аналогично, $|B| = 142 - 14 = 128$ и $|C| = 90 - 9 = 81$. Далее, $A \cap B$ — это

множество всех трёхзначных чисел, делящихся без остатка на 3 и на 7, т. е. кратных числу 21, и аналогичный смысл имеют множества $A \cap C$, $B \cap C$ и $A \cap B \cap C$. Тогда $|A \cap B| = 47 - 4 = 43$, $|A \cap C| = 30 - 3 = 27$, $|B \cap C| = 12 - 1 = 11$, $|A \cap B \cap C| = 4 - 0 = 4$. По формуле включений-исключений $|A \cup B \cup C| = 300 + 128 + 81 - 43 - 27 - 11 + 4 = 432$, а тогда количество трёхзначных чисел, не кратных ни 3, ни 7, ни 11 равно $900 - 432 = 468$.

Ответ: 468 чисел.

Упражнение 38. В летнюю сессию группа студентов из 24 человек сдавала 4 экзамена: физику, высшую математику, экономику и историю России. Те и только те студенты из группы, кто все экзамены сдал с первого раза будут получать стипендию. Известно, что среди студентов группы экономику с первого раза не сдали ровно 8 человек, математику — ровно 10 и физику — ровно 12. При этом 5 человек не смогли с первого раза сдать ни математику ни экономику, 3 — ни экономику, ни физику и 6 — ни экономику ни математику. А ровно двое не смогли сдать ни один из этих трёх предметов. Что касается истории России, то её сдали все студенты с первой же попытки. Сколько же студентов из группы будут получать стипендию в следующем семестре?

3. Математическая индукция

Перед тем как переходить к формулам комбинаторики нам следует познакомиться с одним из важнейших методов доказательства — методом математической индукции. Ранее в этой лекции уже говорилось про этот метод, настало время для его подробного изучения.

На самом деле математическая индукция может быть использована не только при доказательстве, а, например, при построении каких-либо математических объектов, но эти аспекты индукции в настоящей лекции будут опущены, поскольку в ней метод математической индукции нужен исключительно для доказательств тех или иных фактов.

В первом приближении математическая индукция — это обоснование

русской фразы «и так далее по аналогии». Хорошей моделью, демонстрирующей этот метод, является следующая.

Пусть у нас имеется робот стоящий перед достаточно длинной лестницей (математик может считать, что число ступенек лестницы бесконечно; если у кого-то не хватает воображения для такой абстракции, то он может считать, что ступенек просто «очень много»). Зададим роботу следующий алгоритм.

1. Встань на первую ступеньку.

2. Если ты стоишь на ступеньке — перейди на следующую и вернись к этому пункту.

Посмотрим, что произойдёт с роботом при выполнении этого алгоритма. Он встанет на первую ступеньку, затем, согласно пункту 2, перейдёт на вторую, согласно этому же пункту на третью, четвёртую, и т. д. на все последующие, т. е. будет подниматься по лестнице, и пройдёт все её ступеньки (если лестница конечна, то он дойдёт до самого верха, а далее сообщит, что программа невыполнима, поскольку следующей ступеньки нет). Обратите внимание, что мы заставили робота сделать очень много шагов (в идеале бесконечное их количество) всего двумя командами, каждая из которых требовала сделать один шаг. Это и есть индукция.

Заметим также, что при описании действий робота нам пришлось применить слова «и т. д.» В реальной жизни они используются часто, в математике же их применение может привести к грубым ошибкам. Приведём классический пример.

Рассмотрим последовательность $\{a_n\}_{n=1}^{\infty}$ натуральных чисел, определяемую равенством $a_n = n^2 - n + 41$, $n \in \mathbf{N}$. Легко проверить, что $a_1 = 41$, $a_2 = 43$, $a_3 = 47$, $a_4 = 53$ и т. д. Все получаемые числа простые (нацело делятся только на себя и на 1), поэтому рассмотренная последовательность состоит исключительно из простых чисел.

Даже не математику ясно, что аргументация утверждения, мягко говоря, недостаточна: ни из чего не следует, что в последовательности и далее будут идти простые числа. Привести пример, опровергающий утверждение, посложнее, по крайней мере перебор всех натуральных чисел подряд

и проверки результатов на простоту не скоро приведёт к примеру. Но если немного подумать, то можно увидеть, что число $a_{41} = 41^2$ простым не является. Так что утверждение не просто «не доказано», оно в принципе неверно.

Произошло это, как уже говорилось ранее, из-за того, что никак не обоснованы три буквы: сокращение «и т. д.» Однако в примере с роботом мы спокойно использовали это же сокращение и каждому совершенно ясно, что никакой ошибки мы не совершили: доказательство того, что робот пройдёт всю лестницу целиком абсолютно корректно. Обоснование этому содержится в следующей теореме, основной теореме данного пункта — теореме об индукции.

Теорема 15 (об индукции). Пусть для любого натурального числа n сформулировано некое утверждение T_n . При этом известно, что:

- 1) утверждение T_1 верно;
- 2) если для некоторого натурального числа k утверждение T_k верно, то также верно и утверждение T_{k+1} .

Тогда все утверждения T_n верны.

При доказательстве этой теоремы будет использоваться факт, что всякое непустое подмножество натурального ряда содержит наименьший элемент; этот факт (или равносильный ему) является одной из аксиом натурального ряда.

Доказательство теоремы 15. Предположим противное. Тогда множество $A = \{n \in \mathbf{N} : T_n \text{ не верно}\}$ не пусто, поэтому оно имеет наименьший элемент m . В силу первого условия $m \neq 1$, поэтому число $m - 1$ также натурально, и не принадлежит множеству A в силу минимальности m . Значит, T_{m-1} верно, и по второму из условий верно также и T_m . Противоречие, поскольку $m \in A$. Теорема доказана.

Таким образом, для доказательства всей бесконечной последовательно-

сти утверждений достаточно доказать первое утверждение и проверить, что для любого фиксированного натурального числа n из справедливости n -го утверждения следует справедливость $n+1$ -го. Заметим, что при этом доказывать справедливость n -го утверждения не требуется. Такой способ доказательства называется *доказательством по индукции*; при этом доказательство утверждения T_1 называется *базой индукции*, а доказательство истинности утверждения T_{n+1} (в предположении, что T_n верно) — *шагом индукции*, или *индуктивным переходом*. Утверждение T_n называется при этом *индуктивным предположением*.

В качестве примера докажем, что мощность множества всех подмножеств n -элементного множества имеет мощность 2^n .

Пример 21. Доказать, что $|A| = n \Rightarrow |2^A| = 2^n$.

Доказательство. Пусть утверждение T_n таково: для всякого n -элементного множества A мощность множества всех его подмножеств равна 2^n .

База индукции: При $n = 1$, $A = \{a\}$, $2^A = \{\emptyset, A\}$ и утверждение очевидно.

Шаг индукции: Пусть утверждение справедливо для всякого множества мощности k (это — индуктивное предположение). Рассмотрим произвольное множество A мощности $k+1$, т. е. $A = \{a_1, a_2, \dots, a_{k+1}\}$. Пусть $M = \{X \in 2^A : a_{k+1} \notin X\}$. Тогда $M = 2^{\{a_1, a_2, \dots, a_k\}}$ и по индуктивному предположению $|M| = 2^k$. Отображение $f : M \rightarrow 2^A \setminus M$, действующее по правилу $f(X) = X \cup \{a_{k+1}\}$, является биекцией (докажите!), поэтому $|2^A \setminus M| = |M| = 2^k$. Так как множества $2^A \setminus M$ и M не пересекаются, то $|A| = |M| + |2^A \setminus M| = 2^k + 2^k = 2^{k+1}$.

По индукции теорема доказана.

Утверждения T_n обычно в явном виде не формулируют, поскольку чаще всего из контекста ясно, что они собой представляют. Кроме того, теорема об индукции допускает ряд модификаций. Некоторые из них приведены ни-

же без доказательств; желающие легко их докажут самостоятельно, слегка видоизменив приведённое доказательство теоремы 15. После каждой модификации будет приведён пример, показывающий, как она работает.

Теорема 15а. Пусть для любого натурального числа n сформулировано некоторое утверждение T_n , s — натуральное число. При этом известно, что:

- 1) утверждение T_s верно;
- 2) если для некоторого натурального числа k утверждение T_k верно, то также верно и утверждение T_{k+1} .

Тогда все утверждения T_n при $n \geq s$ верны.

Пример 22. Решить в натуральных числах неравенство $2^n > n^2$.

Решение. Очевидно, неравенство выполнено при $n = 1$ и $n = 5$ и не выполнено при $n = 2$, $n = 3$ и $n = 4$. Докажем (применив теорему 15а), что при $n \geq 5$ неравенство выполнено. Пусть утверждение T_n ($n \in \mathbf{N}$) означает, что $2^n > n^2$. Положим $s = 5$. Тогда T_s верно (мы это уже проверили). Осуществим индуктивный переход. Пусть $k > 5$ и $2^k > k^2$. Тогда $2^{k+1} = 2 \cdot 2^k > 2k^2 = k^2 + k^2 > k^2 + 2k + 1 = (k+1)^2$ (так как $k > 5$). Шаг индукции осуществлён; при $n > 5$ неравенство верно.

Ответ: $n \in \mathbf{N} \setminus \{2, 3, 4\}$

Теорема 15б. Пусть для любого натурального числа n сформулировано некоторое утверждение T_n , s — натуральное число. При этом известно, что:

- 1) утверждения T_1, T_2, \dots, T_s верны;
- 2) если для некоторого натурального числа $k > s$ верны все утверждения $T_{k-s}, T_{k-s+1}, \dots, T_{k-1}$, то также верно и утверждение T_k .

Тогда все утверждения T_n верны.

Пример 23. Последовательность a_n задана рекуррентно: $a_1 = 1$,

$a_2 = 2$, $a_n = a_{n-1} + 2a_{n-2}$. Доказать, что $a_n = 2^{n-1}$.

Решение. Применим теорему 15б для $s = 2$. База индукции очевидна: $a_1 = 2^{1-1} = 1$, $a_2 = 2^{2-1} = 2$. Пусть k таково, что $a_{k-2} = 2^{k-3}$ и $a_{k-1} = 2^{k-2}$. Тогда $a_k = a_{k-1} + 2a_{k-2} = 2^{k-2} + 2 \cdot 2^{k-3} = 2^{k-1}$. Шаг индукции осуществлён, утверждение доказано.

Теорема 15в. Пусть для любого натурального числа n сформулировано некое утверждение T_n . При этом известно, что:

- 1) утверждение T_1 верно;
- 2) если для некоторого натурального числа k верны все утверждения T_1, T_2, \dots, T_{k-1} , то также верно и утверждение T_k .

Тогда все утверждения T_n при $n \geq s$ верны.

Пример 24. В деревне n домов и m непересекающихся заборов. Известно, что каждый забор ограждает некоторую непустую совокупность домов и что никакие два забора не ограждают одну и ту же совокупность домов. Докажите, что $m \leq 2n - 1$.

Сперва приведём пример неправильного применения индукции к решению этой задачи (заметим, что такую ошибку допускают многие студенты): При $n = 1$ утверждение, очевидно, верно: один дом может быть окружён только одним забором — это база индукции. Индуктивное предположение: пусть утверждение верно при $n = k - 1$. Возьмём $k - 1$ дом и построим возможно большее число заборов, удовлетворяющих условию задачи; их не больше, чем $2k - 3$. Добавим ещё один дом. Теперь мы можем достроить максимум 2 забора: один непосредственно вокруг добавленного дома и один вокруг всех k домов. Всего заборов будет не больше, чем $2k - 3 + 2 = 2k - 1$. Утверждение доказано.

Где же здесь ошибка? Она заключается в том, что далеко не всякую совокупность домов и заборов можно получить способом, рассмотренном в решении, т. е. добавлением к максимальной совокупности домов и заборов

ещё одного дома и нескольких заборов. Так, уже при 4-х домах возможна ситуация, когда построены заборы, окружающие каждый дом, заборы ограждающие совокупности домов (1, 2) и (3, 4), и забор, ограждающий все 4 дома. Эта конструкция, очевидно, не может быть получена из максимальной конструкции для трёх заборов. Для неё рассуждения индуктивного перехода не работают. Верное же решение такое.

Решение. База индукции очевидна. В качестве индуктивного предположения примем, что при всех $n < k$ в совокупности из n домов нельзя построить более $2n - 1$ забора. Рассмотрим максимальную совокупность заборов при наличии k домов. Имеется внешний забор, ограждающий все дома. Снесём его. Останется ровно две группы домов, каждая ограниченная своим забором (если групп больше, изначальное число заборов не было максимальным, можно было построить ещё один, ограничивающий любые две группы в совокупности). Пусть в этих группах p и s домов ($p + s = k$). Так как группы непусты, $p < k$ и $s < k$, по индуктивному предположению в группах не более, чем $2p - 1$ и $2s - 1$ заборов соответственно. Всего заборов не более, чем 1 (внешний) $+(2p - 1) + (2s - 1) = 2(p + s) - 1 = 2k - 1$. Утверждение доказано.

Упражнение 39. Докажите теоремы 15а, 15б и 15в.

Упражнение 40. Докажите методом математической индукции следующие утверждения

- а) $1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$ при всех натуральных n ;
- б) для всех $n \in \mathbf{N}$, $n > 4$ выполнено неравенство $2^n > 2n + 7$;
- в) число $7^n + 5$ делится без остатка на 6 при всех натуральных n .

В заключении отметим, что в математике существует более общий метод индукции — так называемый *метод трансфинитной индукции*. Рассказ об этом методе выходит за рамки нашего курса; желающие могут ознакомиться с ним в соответствующей литературе.

ЛЕКЦИЯ VII. ОСНОВЫ КОМБИНАТОРИКИ. ОПРЕДЕЛЕНИЯ И ФОРМУЛЫ

1. Размещения, перестановки и сочетания

В комбинаторике (как, впрочем, и в других разделах математики) имеются конструкции, которые встречаются в задачах настолько часто, что для них разумно выделить специальные обозначения, и свойства этих конструкций изучить и запомнить. В комбинаторике это размещения, сочетания и перестановки. Дадим два определения каждому из этих понятий, одно — на языке теории множеств, другое — на языке выбора, и пусть каждый сможет составить собственное мнение о том, каким из двух определений удобнее пользоваться.

Начнём с языка выбора. Давая формулировку на этом языке, мы будем представлять исходное множество в виде некоторой коробочки или мешочка, из которого извлекаются элементы. Будем считать, что элементы извлекаются по одному в количестве m штук (m — натуральное число или ноль), а исходное множество содержит n элементов, $n \in \mathbf{N}$ (не обязательно $m \leq n$). Выбирать элементы можно принципиально двумя способами: 1) выбранные элементы обратно в коробочку не возвращаются (и, следовательно, не могут быть выбраны повторно), тогда говорят, что мы имеем дело с *бесповторной выборкой*; 2) каждый выбранный элемент записывается и возвращается обратно в коробочку, после чего может быть снова выбран на последующих шагах, говорят, что в этом случае наблюдается *выборка с повторениями*. Кроме того, вне зависимости от того, какой способ выбора имеет место, возможны две ситуации: I) порядок выбранных элементов не существен, важен только их набор. В этом случае мы имеем дело с *сочетаниями*; II) Порядок, в каком элементы выбираются, существен. Тогда речь идёт о *размещениях*. Комбинируя римские и арабские числа, можно получить 4 типа выборок: 1I — сочетания без повторений; 2I — сочетания с повторениями; 1II — размещения без повторений; 2II — размещения с повторениями. Кроме того, имеется особый случай: бесповторная выборка всех элементов исходного множества — она называется *перестанов-*

кой, поскольку может трактоваться следующим образом: упорядочиваются все элементы исходного множества в соответствии с порядком их извлечения (первому извлечённому элементу присваивают номер 1, второму — номер 2, и так до последнего, которому присваивают номер n). Получаем определения на языке выбора, а именно.

Определение 23. Пусть дано непустое конечное множество T , $|T| = n$, и пусть $m \in \mathbf{N} \cup \{0\}$. Тогда:

а) *размещением с повторениями по m множества T* называют любой способ выбрать из множества T m элементов, при условии, что порядок выбираемых элементов существенен, и имеет место выборка с повторениями;

б) *размещением без повторений по m множества T* называют любой способ выбрать из множества T m элементов, при условии, что порядок выбираемых элементов существенен, и имеет место бесповторная выборка;

в) *перестановкой* множества T называют любой способ упорядочить множество T ;

г) *сочетанием без повторений по m множества T* называют любой способ выбрать из множества T m элементов, при условии, что порядок выбираемых элементов не важен, и имеет место бесповторная выборка;

д) *сочетанием с повторениями по m множества T* называют любой способ выбрать из множества T m элементов, при условии, что порядок выбираемых элементов не важен, и имеет место выборка с повторениями.

На языке теории множеств определение 23 выглядит следующим образом. Вам предоставляется возможность самостоятельно убедиться, что одноимённые пункты определений 23 и 23' говорят об одном и том же объекте.

Определение 23'. Пусть дано непустое конечное множество T , $|T| = n$, и пусть $m \in \mathbf{N} \cup \{0\}$. Тогда:

а) *размещением с повторениями по m множества T* называют любой

элемент m -ой декартовой степени множества T (т. е. любой элемент множества T^m);

б) *размещением без повторений по m* множества T называют любой элемент m -ой декартовой степени множества T (т. е. любой элемент множества T^m), координаты которого попарно различны;

в) *перестановкой* множества T называют любое размещение без повторений по n множества T ;

г) *сочетанием без повторений по m* множества T называют любое подмножество множества T мощности m .

Заметим, что в определении 23' отсутствует понятие сочетания с повторениями. Связано это с тем, что теоретико-множественный аналог этого понятия достаточно сложен. По сути повторная выборка — это множество, в которое элементы могут входить не по одному разу. Но с точки зрения теории множеств (см. лекцию 1) считается, что всякий элемент во множество входит ровно один раз, т. е. мы не различаем, к примеру, множества $\{a, a, b\}$ и $\{a, b, b\}$. А говоря о сочетании с повторениями, такие множества необходимо различать, чтобы не вводить каких-нибудь новых понятий и не усложнять этим материал курса, мы просто не будем давать теоретико-множественного определения сочетанию с повторениями, ограничившись языком выбора. Кроме этого можно заметить, что в определении 23' в пунктах (а) и (б) при $m = 0$ возникают понятия нулевой декартовой степени множества, в то время как декартова степень T^n была определена только для натуральных чисел n . Этот недостаток ликвидируется, если доопределить понятие размещения (и с повторениями, и без) для случая $m = 0$. Делается это просто, а именно, считается, что для всякого множества T *единственным размещением по 0* является пустое множество. Что до пунктов (в) и (г) определения 23', то там такой проблемы нет.

Чаще всего, когда речь идёт о сочетаниях без повторений или о размещениях без повторений, слова «без повторений» опускают и говорят просто о сочетаниях и размещениях. Когда речь идёт о тех же понятиях, но с повторениями, об этом всегда говорится явно. Мы не будем в нашем кур-

се вводить понятие перестановки с повторениями, хотя в ряде учебников по комбинаторике оно приводится. Желающие могут ознакомиться с ним самостоятельно.

Количество всех сочетаний (без повторений) n -элементного множества по m обозначается символом C_n^m , сочетаний с повторениями — символом \overline{C}_n^m , размещений (без повторений) — символом A_n^m , размещений с повторениями по m n -элементного множества обозначается символом \overline{A}_n^m , наконец, количество всех перестановок n -элементного множества записывается как P_n . Наша ближайшая цель получить явные формулы для всех этих чисел (как функции от m и n). Эти формулы будут сформулированы и доказаны как теоремы; после каждой такой теоремы будет приведён простой пример на использование соответствующего понятия и соответствующей формулы.

Начнём с размещений с повторениями, поскольку это самое простое.

Теорема 16. Для любых чисел $n \in \mathbf{N}$ и $m \in \mathbf{N} \cup \{0\}$ верно равенство $\overline{A}_n^m = n^m$.

Доказательство. Пусть T — произвольное множество, содержащее n элементов. Тогда, согласно определению 23' (а), нам требуется найти количество элементов во множестве T^m . По теореме 13 это количество равно $|T|^m$ для любого натурального числа m . Для завершения доказательства остаётся рассмотреть случай $m = 0$. Но единственное размещение по 0 — это пустое множество, поэтому $|A_n^0| = |\{\emptyset\}| = 1 = n^0$, поэтому доказываемая формула верна и в этом случае.

Пример 25. Андрей, Борис, Виктор, Григорий и Дмитрий сдают экзамен по математике. Оценка за экзамен может быть либо «отлично», либо «хорошо», либо «удовлетворительно», либо «неудовлетворительно». Сколькими способами могут распределиться оценки?

Решение. Нам требуется из четырёх возможных оценок выбрать 5, причём порядок выбора важен (ведь есть разница в том, например, кто из

студентов, Виктор или Дмитрий, сдаст на «хорошо», а кто на «неудовлетворительно»), а повторения допускаются. Значит, мы имеем дело с размещением с повторениями, и количество способов поставить оценки равно \overline{A}_4^5 , что по теореме 15 равно 4^5 .

Ответ: 4^5 способов.

Теперь поговорим о размещениях без повторений. Пусть, по-прежнему, имеется некоторое n -элементное множество T , $n \in \mathbf{N}$, и нам требуется выбрать из него m различных элементов с учётом их порядка. Прежде всего заметим, что при $m > n$ это просто неосуществимо, так как ни из какого множества нельзя выбрать больше элементов, чем это множество содержит. Таким образом, в этом случае $A_n^m = 0$. Далее будем полагать $0 \leq m \leq n$. Случаи $m = 0$ и $m = 1$ тривиальны: существует ровно один способ выбрать 0 элементов (не выбирать ничего), и ровно n способов выбрать один элемент (любой элемент исходного множества может быть выбран). Пусть $1 < m \leq n$. Тогда первый элемент размещения может быть выбран n способами. Легко заметить, что при любом способе выбора первого элемента второй элемент размещения может быть выбран $n - 1$ числом способов, независимо от этого выбора третий элемент может быть выбран $n - 2$ способами, и так вплоть до m -го элемента, который может быть выбран $n - (m - 1) = n - m + 1$ числом способов (строгое доказательство может быть легко проведено методом математической индукции). Тогда по теореме 12 количество способов осуществить требуемый выбор равно

$$\begin{aligned} & n \cdot (n - 1) \cdot \dots \cdot (n - m + 1) = \\ &= \frac{n \cdot (n - 1) \cdot \dots \cdot (n - m + 1) \cdot (n - m) \cdot (n - m - 1) \cdot 2 \cdot 1}{(n - m) \cdot (n - m - 1) \cdot 2 \cdot 1} = \frac{n!}{(n - m)!}. \end{aligned}$$

(Напомним, что для натурального числа k через $k!$ обозначается произведение всех натуральных чисел, не превосходящих числа k .) Ещё заметим, что полученная формула верна также и при $m = 0$ $\left(\frac{n!}{(n - 0)!} = 1\right)$, и при $m = 1$ $\left(\frac{n!}{(n - 1)!} = n\right)$, а при $m > n$ она не применима: в знаменателе

стоит неопределённое выражение — факториал отрицательного числа. Мы доказали следующую теорему.

Теорема 17. Для любых чисел $n \in \mathbf{N}$ и $m \in \mathbf{N} \cup \{0\}$ верно равенство

$$A_n^m = \begin{cases} 0, & \text{при } m > n \\ \frac{n!}{(n-m)!} & \text{при } 0 \leq m \leq n \end{cases}.$$

Замечание. При $m \neq 0$ число размещений без повторений может быть вычислено по формуле $A_n^m = n \cdot (n-1) \cdot \dots \cdot (n-m+1)$, т. е. как произведение m последовательных целых чисел, наибольшее из которых равно n . Эта последняя формула останется верной и при $m > n$, так как в этом случае среди множителей возникает число 0.

Пример 26. В группе обучается 18 студентов. Каждый день один из студентов назначается дежурным по аудитории. Сколькими способами может быть составлено расписание дежурств на неделю (с понедельника по субботу), если ни один из студентов не должен дежурить более одного раза?

Решение. Нам требуется из группы в 18 человек выбрать шестерых дежурных. Повторы исключены, так как каждый студент дежурит не более одного дня. Порядок важен, так как первый студент будет дежурить именно в понедельник, а не в другой день недели, второй — исключительно во вторник и т. д. Значит, мы имеем дело с размещением без повторений. Искомое число равно $A_{18}^6 = \frac{18!}{12!} = 18 \cdot 17 \cdot \dots \cdot 13$.

Ответ: $\frac{18!}{12!}$ способами.

Поскольку перестановки — частный случай размещения без повторений в качестве следствия теоремы 17 получаем важную формулу.

Теорема 18. Для любого $n \in \mathbf{N}$ верно равенство $P_n = n!$

Доказательство. $P_n = A_n^n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = \frac{n!}{1} = n!$.

Упражнение 41. Докажите теорему 18 не используя формулу числа перестановок, методом математической индукции.

Теперь давайте на минуту откажемся от того, что число n — натуральное, и применим формулу теоремы 18 для случая $n = 0$. Получим, что число способов упорядочить множество, содержащее 0 элементов (а мы уже знаем, что такое множество является пустым), равно $0!$, т. е. равно 1. Это вполне соответствует здравому смыслу: действительно пустое множество можно упорядочить ровно одним способом: ничего не упорядочивать. Это рассуждение объясняет, почему математики положили $0!$ равным 1, а не нулю (вопрос, интересующий всех, кто впервые сталкивается с понятием факториала).

Ввиду простоты теоремы 18, пример, её иллюстрирующий, подберём сложнее.

Пример 27. На званый обед приглашены 10 человек: 5 мужчин и 5 женщин. Сколькими различными способами их можно рассадить за круглый стол, если по правилам этикета за столом мужчины и женщины должны чередоваться? При этом нас интересует только порядок приглашённых за столом друг относительно друга, т. е. способы, получающиеся пересадкой всех участников обеда на одно и то же число мест по часовой стрелке, считаются одинаковыми.

Решение. Сначала определим для каждой приглашённой женщины мужчину, который будет сидеть от неё, скажем, по правую руку. Для этого упорядочим всех приглашённых женщин произвольным образом (например, по знатности или по росту) и пусть согласно этому порядку женщины последовательно выбирают себе правого соседа из числа ещё не занятых мужчин. Тогда распределение по парам может быть осуществлено $P_5 = 5! = 120$ числом способов. Теперь упорядочим полученные пары, это можно сделать также $5! = 120$ способами. По теореме об умножении всего возможны $5! \cdot 5!$ вариантов. Теперь первую пару сажаем за стол произволь-

но (женщина слева от мужчины), вторую — вплотную к ней слева, третью — слева от второй и т. д. вплоть до последней пятой пары. Таким образом мы получим всевозможные допустимые рассадки. Однако при этом каждая рассадка будет получена пятью способами — в зависимости от того, какую из 5 пар мы посадим первой (т. е. упорядочения пар, получаемые циклическим сдвигом нам дают одну и ту же рассадку). Значит, всего способов рассадить требуемым образом приглашённых будет в 5 раз меньше, т. е. число этих способов будет равно $\frac{5! \cdot 5!}{5} = 5! \cdot 4! = 2880$.

Ответ: 2880 способами.

Упражнение 42. Прямоугольный лист клетчатой бумаги имеет размеры 10 клеток в ширину и 15 клеток в длину. Клетки этого листа раскрашиваются в 10 цветов (каждая клетка в свой цвет) так, чтобы в каждом 10-клеточном столбце встречались все 10 цветов. Сколькими способами может быть осуществлена такая раскраска?

Перейдём к сочетаниям. Сначала к сочетаниям без повторений.

Теорема 19. Для любых чисел $n \in \mathbf{N}$ и $m \in \mathbf{N} \cup \{0\}$ верно равенство

$$C_n^m = \begin{cases} 0, & \text{при } m > n \\ \frac{n!}{m! \cdot (n-m)!} & \text{при } 0 \leq m \leq n \end{cases}.$$

Доказательство. Не вызывает сомнения верхняя строчка в формуле: при $m > n$ не существует ни одного сочетания, поскольку из n -элементного множества невозможно без повторения выбрать более, чем n элементов. Пусть теперь $m \leq n$. Подсчитаем число всевозможных размещений (без повторений) по m . С одной стороны, их A_n^m . С другой стороны, каждое размещение по m может быть получено следующим образом: берём некоторое сочетание по m и произвольно переставляем его элементы. Всего сочетаний по m ровно C_n^m , упорядочить m -элементное множество можно P_m способами, поэтому общее число размещений по m равно (согласно теореме 13) $C_n^m \cdot P_m$. Из равенства $A_n^m = C_n^m \cdot P_m$ находим, что

$C_n^m = \frac{A_n^m}{P_m^m}$. Остаётся подставить в последнюю формулу выражения для A_n^m и P_m^m , полученные в теоремах 17 и 18.

Замечание. Подсчёт числа сочетаний непосредственно по формуле, т. е. через вычисления факториалов, очень громоздок. Это связано с тем, что число $n!$ очень быстро возрастает с ростом n . Так, уже $10!$ равно 3 628 800, а далее при каждом возрастании числа n на 1 десятичная запись числа $n!$ удлиняется как минимум на один разряд. А число $60!$ уже превосходит 10^{100} и становится недоступным и стандартному школьному калькулятору. Поэтому если вам понадобится получить десятичную запись числа вида C_n^m , ищите другие способы его подсчёта.

Один из таких способов заключается в следующем: для подсчёта числа C_n^m сперва записывается дробная черта. Затем в знаменателе выписывают первые m натуральных чисел, чередуя их знаком умножения, а в числителе выписывают в качестве множителей m последовательных натуральных чисел, наибольшее из которых равно n . Далее производят многочисленые сокращения (при этом знаменатель сокращается до 1) и только затем перемножают оставшиеся в числителе числа. Вот как выглядит такой счёт для $n = 12$, $m = 4$:

$$C_{12}^4 = \frac{12 \cdot 11 \cdot 10 \cdot 9}{1 \cdot 2 \cdot 3 \cdot 4} = \frac{11 \cdot 10 \cdot 9}{2} = 11 \cdot 5 \cdot 9 = 495.$$

В следующей части лекции подробно разберём некоторые свойства чисел вида C_n^m и станет понятно, как можно быстро их находить. Но сначала пара примеров, иллюстрирующих теорему 19.

Пример 28. В караульном взводе имеется 3 офицера, 6 сержантов и 20 рядовых. Из взвода требуется выделить караул в составе одного офицера, двух сержантов и 6 рядовых. Сколькими способами можно это сделать?

Решение. Офицера надо выбрать одного из трёх, и ясно, что это можно сделать тремя способами — по числу офицеров во взводе. Сержантов надо выбрать двух, и осуществить сей выбор можно C_6^2 способами. Наконец, 6 рядовых можно выбрать C_{20}^6 способами. По теореме об умножении получаем ответ.

Ответ: $3 \cdot C_6^{2 \cdot 6}{}_{20}$ способами.

Пример 29. Четверо пенсионеров играют в домино. В начале каждого кона все 28 косточек домино раздаются (случайным образом) игрокам (каждый получает по 7 косточек). Сколькими способами может быть осуществлена раздача?

Решение. Если мы будем раздавать косточки пенсионерам по одной (как при раздаче в карты), то подсчёт числа способов окажется сложноватым. Гораздо проще считать, что сначала даётся 7 косточек одному из пенсионеров (назовём его пенсионером А), затем 7 — другому (пенсионеру Б), затем 7 — пенсионеру В, а оставшиеся 7 — пенсионеру Г. Тогда пенсионеру А косточки могут быть выданы C_{28}^7 способами (не вздумайте находить десятичную запись этого числа — наверняка ошибётесь в счёте). Пенсионеру Б будут выделены также 7 косточек, но уже из числа 21 оставшейся после того, как пенсионер А свои косточки получил — это C_{21}^7 способов. Аналогично, пенсионер В может получить свои косточки C_{14}^7 способами, а пенсионер Г — только одним (формально C_7^7) способом. По теореме об умножении общее число способов раздать домино равно $C_{28}^7 \cdot C_{21}^7 \cdot C_{14}^7 \cdot 1$. Ответ можно оставить и в таком виде, но лучше преобразовать его так:

$$C_{28}^7 \cdot C_{21}^7 \cdot C_{14}^7 \cdot 1 = \frac{28!}{7! \cdot 21!} \cdot \frac{21!}{7! \cdot 14!} \cdot \frac{14!}{7! \cdot 7!} = \frac{28!}{7! \cdot 7! \cdot 7! \cdot 7!} = \frac{28!}{(7!)^4}.$$

Ответ: $\frac{28!}{(7!)^4}$ способами.

Упражнение 43. В хоккейной команде 22 игрока: 2 вратаря, а остальные — полевые игроки, разбитые на четыре пятёрки. Требуется выбрать из них капитана команды и трёх его заместителей. Сколькими способами можно осуществить такой выбор? Как изменится ответ, если потребовать, чтобы в каждой из пятёрок был либо капитан, либо его заместитель?

Пример 30. Пусть даны натуральные числа n_1, n_2, \dots, n_k . Докажите, что число $\frac{(n_1 + n_2 + \dots + n_k)!}{n_1! \cdot n_2! \cdot \dots \cdot n_k!}$ является целым.

Доказательство. Рассмотрим множество, состоящее из $n_1 + n_2 + n_3 + \dots + n_k$ различных предметов (например, можете считать, что это — книги разных авторов) и k человек, которым надо эти предметы раздать. При этом i -ый человек ($i = \overline{1, k}$) должен получить ровно n_i предметов. Найдём число способов l осуществить требуемую раздачу — ясно, что l — число натуральное. Рассуждая как в примере 29, получаем, что

$$\begin{aligned} l &= C_{n_1+n_2+\dots+n_k}^{n_1} \cdot C_{n_2+\dots+n_k}^{n_2} \cdot \dots \cdot C_{n_{k-1}+n_k}^{n_{k-1}} = \\ &= \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! \cdot (n_2 + \dots + n_k)!} \cdot \frac{(n_2 + \dots + n_k)!}{n_2! \cdot (n_3 + \dots + n_k)!} \cdot \dots \cdot \frac{(n_{k-1} + n_k)!}{n_{k-1}! \cdot n_k!} = \\ &= \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! \cdot n_2! \cdot \dots \cdot n_k!}. \end{aligned}$$

Итак, данное в условии число и есть число l . Утверждение доказано.

Упражнение 44. Докажите, что для любого натурального числа n произведение любых последовательных n натуральных чисел делится без остатка на $n!$.

У нас остались пока без внимания сочетания с повторением. Но перед тем, как перейти к ним мы поподробнее изучим числа вида C_n^m , поскольку оказалось, что они играют серьёзную роль не только в комбинаторике, но и вообще в математике.

2. Биномиальные коэффициенты и их свойства

Числа вида C_n^m , где $m \leq n$, $m \in \mathbf{N} \cup \{0\}$, $n \in \mathbf{N}$ часто называют *биномиальными коэффициентами*. Чуть позже мы поймём, с чем связано это название. Пока же просто будем изучать, как устроены такие числа. Свойства биномиальных коэффициентов мы будем выписывать в виде утверждений, весьма простых, почти очевидных. Для каждого такого утверждения будет приведено два доказательства — одно чисто алгебраическое, другое — комбинаторное. Каждый может выбрать то доказательство, какое ему больше понравится или будет более понятным. Начнём с простого.

Утверждение 1. $C_n^0 = C_n^n = 1$ при всех натуральных n .

Доказательство первое (комбинаторное). C_n^0 — это количество способов выбрать 0 элементов из n имеющихся. Ясно, что такой способ только один: не выбирать ничего. Аналогично, C_n^n — это количество способов выбрать n элементов из n . Такой способ тоже единственный: выбрать всё.

Доказательство второе (алгебраическое). Подставим поочерёдно $m = 0$ и $m = n$ в формулу числа сочетаний $C_n^m = \frac{n!}{m! \cdot (n - m)!}$:

$$C_n^0 = \frac{n!}{0! \cdot (n - 0)!} = \frac{n!}{1 \cdot n!} = 1; \quad C_n^n = \frac{n!}{n! \cdot (n - n)!} = \frac{1}{0!} = 1.$$

Утверждение 2. $C_n^1 = n$ при всех натуральных n .

Доказательство первое (комбинаторное). Нас спрашивают количество способов, каким можно выбрать 1 элемент из n имеющихся. Но таких способов ровно n : можно выбрать первый элемент, можно — второй и т. д. вплоть до n -го.

Доказательство второе (алгебраическое). Подставим в формулу числа сочетаний $m = 1$:

$$C_n^1 = \frac{n!}{1! \cdot (n - 1)!} = \frac{1 \cdot 2 \cdot \dots \cdot (n - 1) \cdot n}{1 \cdot 2 \cdot \dots \cdot (n - 1)} = n.$$

Утверждение 3. $C_n^m = C_n^{n-m}$ при всех натуральных n и $m \leq n$.

Доказательство первое (комбинаторное). Выбрать m элементов из n имеющихся — это то же самое, что указать те $n - m$ элементов, которые не будут выбраны. Таким образом, налицо биекция между множеством способов выбрать m элементов и множеством способов выбрать $n - m$ элементов, поэтому таких способов поровну, о чём и говорит доказываемая формула.

Доказательство второе (алгебраическое). Проведём несложные

алгебраические выкладки:

$$C_n^{n-m} = \frac{n!}{(n-m)! \cdot (n-(n-m))!} = \frac{n!}{(n-m)! \cdot m!} = C_n^m.$$

Утверждение 4. $C_n^m + C_n^{m+1} = C_{n+1}^{m+1}$ при всех натуральных n и $m \leq n-1$.

Доказательство первое (комбинаторное). Рассмотрим какое-нибудь множество, содержащее $n+1$ элемент и отметим один из его элементов. C_{n+1}^{m+1} — это количество $(m+1)$ -элементных подмножеств этого множества. Среди этих подмножеств есть такие, которые не содержат отмеченный элемент — тогда они содержат $m+1$ неотмеченный, и ясно, что таких подмножеств C_n^{m+1} — подойдёт любое $(m+1)$ -элементное подмножество множества неотмеченных элементов. Остальные подмножества содержат отмеченный элемент, следовательно, содержат ровно m неотмеченных. Значит, таких подмножеств C_n^m . По теореме о сложении имеем $C_n^m + C_n^{m+1} = C_{n+1}^{m+1}$, ч. т. д.

Доказательство второе (алгебраическое). Здесь выкладки посложнее:

$$\begin{aligned} C_n^m + C_n^{m+1} &= \frac{n!}{m!(n-m)!} + \frac{n!}{(m+1)!(n-m-1)!} = \\ &= \frac{n!}{m!(n-m-1)!(n-m)} + \frac{n!}{(m)!(m+1)(n-m-1)!} = \\ &= \frac{n!(m+1) + n!(n-m)}{m!(m+1)(n-m-1)!(n-m)} = \frac{n!(m+1+n-m)}{(m+1)!(n-m)!} = \\ &= \frac{(n+1)!}{(m+1)!((n+1)-(m+1))!} = C_{n+1}^{m+1}. \end{aligned}$$

Утверждение 5. $C_n^0 + C_n^1 + \dots + C_n^n = 2^n$ при всех натуральных n .

Доказательство первое (комбинаторное). Зафиксируем n -элементное множество T . Каждое слагаемое вида C_n^k ($0 \leq k \leq n$) означает количество k -элементных подмножеств множества T ; значит, в левой части формулы стоит количество всех подмножеств множества T , содержащих не более n элементов. Но такими являются все подмножества T ,

и как было показано при изучении индукции (пример 21), их число равно 2^n . Утверждение доказано.

Доказательство второе (алгебраическое). Докажем утверждение индукцией по n .

База индукции. Пусть $n = 1$, тогда $C_1^0 + C_1^1 = 1 + 1 = 2 = 2^1$ — утверждение верно.

Шаг индукции. Пусть утверждение верно при некотором натуральном числе n , т. е. пусть $C_n^0 + C_n^1 + \dots + C_n^n = 2^n$. Докажем, что $C_{n+1}^0 + C_{n+1}^1 + \dots + C_{n+1}^n + C_{n+1}^{n+1} = 2^{n+1}$. Действительно, применяя $n - 1$ раз утверждение 4, получим

$$\begin{aligned} & C_{n+1}^0 + C_{n+1}^1 + C_{n+1}^2 + \dots + C_{n+1}^n + C_{n+1}^{n+1} = \\ &= C_{n+1}^0 + C_n^0 + C_n^1 + C_n^1 + C_n^2 + \dots + C_n^{n-1} + C_n^n + C_{n+1}^{n+1} = \\ &= (C_{n+1}^0 + C_n^0) + 2C_n^1 + \dots + 2C_n^{n-1} + (C_n^n + C_{n+1}^{n+1}) = \\ &= 2(C_n^0 + C_n^1 + \dots + C_n^n) = 2 \cdot 2^n = 2^{n+1}. \end{aligned}$$

(Здесь учтено, что $C_{n+1}^0 = 1 = C_n^0$ и что $C_n^n = 1 = C_{n+1}^{n+1}$.) Шаг индукции, а с ним и всё утверждение 5 доказаны.

3. Бином Ньютона и треугольник Паскаля

Следующее свойство чисел C_n^m очень известно и настолько важно в математике, что мы назовём его не утверждением, а теоремой. Оно связано с понятием *бинома Ньютона* и связано настолько сильно, что числа C_n^m часто называют *биномиальными коэффициентами* — из формулировки теоремы станет понятным, откуда взялось это название. Биномом Ньютона в математике называют выражение $(a+b)^n$ (где n — натуральное число, а a и b могут быть как числами, так и переменными), а также то выражение, которое получается из $(a+b)^n$ после раскрытия скобок. Само слово «бином» в переводе на русский язык означает «двучлен»; легко заметить, что бином Ньютона двучленом, вообще говоря, не является: после раскрытия скобок слагаемых получится больше, чем два. Но словосочетание «бином

Ньютона» как-то прижилось в математике, причём зачастую под ним понимают не само выражение, а правило раскрытия скобок в этом самом биноме. Нам всё же представляется правильным отнести это словосочетание к выражению $(a + b)^n$ (пусть это и многочлен более высокой степени, чем два), а указанное правило назвать *теоремой о биноме Ньютона*.

Теорема 20 (о биноме Ньютона). Для любых чисел a и b и любого натурального числа n справедлива формула

$$\begin{aligned}(a + b)^n &= \sum_{i=0}^n C_n^i a^{n-i} b^i = \\ &= a^n + \frac{n}{1} a^{n-1} b + \frac{n(n-1)}{1 \cdot 2} a^{n-2} b^2 + \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} a^{n-3} b^3 + \dots + b^n.\end{aligned}$$

Доказательство. Как и при доказательстве остальных свойств чисел сочетаний, приведём два способа: комбинаторный и алгебраический.

Комбинаторный способ. Представим бином в виде

$$(a + b)^n = \underbrace{(a + b) \cdot (a + b) \cdot \dots \cdot (a + b)}_{n \text{ скобок}}. \quad (*)$$

По правилу умножения многочленов $(*)$ представляет из себя сумму всевозможных произведений вида $x_1 \cdot x_2 \cdot \dots \cdot x_n$, где каждое x_i — одно из слагаемых в i -ой скобке, т. е. $x_i \in \{a, b\}$ для всех $1 \leq i \leq n$. Если в таком произведении число a встречается k раз ($0 \leq k \leq n$), то число b встретится в нём $n - k$ раз, значит, всё произведение будет равно $a^k \cdot b^{n-k}$. Теперь подсчитаем сколько раз встретится выражение $a^k \cdot b^{n-k}$. Надо выбрать $n - k$ скобок в произведении $(*)$, из которых будем брать число b (из остальных скобок берём число a). Это можно сделать C_n^{n-k} числом способов — столько слагаемых вида $a^k \cdot b^{n-k}$ и получится. Таким образом, имеет место формула $(a + b)^n = \sum_{i=0}^n C_n^i a^{n-i} b^i$, ч. т. д.

Алгебраический способ. Докажем равенство индукцией по n . База индукции очевидна: $(a + b)^1 = a + b = C_1^0 a^1 b^0 + C_1^1 a^0 b^1$. Осуществим шаг индукции. Пусть для некоторого натурального n выполнено равенство $(a + b)^n = \sum_{i=0}^n C_n^i a^{n-i} b^i$ — это индуктивное предположение. Надо доказать,

что $(a + b)^{n+1} = \sum_{i=0}^{n+1} C_{n+1}^i a^{n+1-i} b^i$. Проведем преобразования:

$$\begin{aligned}
(a + b)^{n+1} &= (a + b) \cdot (a + b)^n = (a + b) \cdot \sum_{i=0}^n C_n^i a^{n-i} b^i = \\
&= \sum_{i=0}^n C_n^i a^{n+1-i} b^i + \sum_{i=0}^n C_n^i a^{n-i} b^{i+1} = \sum_{i=0}^n C_n^i a^{n+1-i} b^i + \sum_{i=1}^{n+1} C_n^{i-1} a^{n-(i-1)} b^i = \\
&= C_n^0 a^{n+1} b^0 + \sum_{i=1}^n C_n^i a^{n+1-i} b^i + \sum_{i=1}^n C_n^{i-1} a^{n-(i-1)} b^i + C_n^n a^0 b^{n+1} = \\
&= C_{n+1}^0 a^{n+1} b^0 + \sum_{i=1}^n (C_n^i + C_n^{i-1}) a^{n+1-i} b^i + C_{n+1}^{n+1} a^0 b^{n+1} = \\
&= C_{n+1}^0 a^{n+1} b^0 + \sum_{i=1}^n C_{n+1}^i a^{n+1-i} b^i + C_{n+1}^{n+1} a^0 b^{n+1} = \sum_{i=0}^{n+1} C_{n+1}^i a^{n+1-i} b^i.
\end{aligned}$$

Чтобы лучше понять доказательство, опишем словами каждое из равенств приведённой цепочки. Первое равенство понятно: отщеплён один из множителей, чтобы можно было применить индуктивное предположение. Второе равенство и есть его применение. Следующий переход — это простое раскрытие скобок. Наиболее неочевидным является переход от $\sum_{i=0}^n C_n^i a^{n-i} b^{i+1}$ к $\sum_{i=1}^{n+1} C_n^{i-1} a^{n-(i-1)} b^i$. На самом деле, здесь мы просто заменили i на другую переменную $i - 1$, что допустимо, поскольку переменная i — есть индекс суммирования. Разумеется, пределы суммирования поменялись: если $0 \leq i - 1 \leq n$, то $1 \leq i \leq n + 1$. В следующем равенстве отцепили от каждой суммы по слагаемому: от первой суммы — слагаемое с индексом 0, от второй — с индексом $n + 1$, чтобы можно было осуществить сложение. В следующем равенстве просто приведены подобные, а далее надо воспользоваться утверждением 4 и фактами, что $C_n^0 = C^0 n + 1$, а $C_n^n = C_{n+1}^{n+1}$. Наконец, в последнем действии собраны три слагаемых под один знак суммы, при этом расширены границы суммирования. Всем этим равенство $(a + b)^{n+1} = \sum_{i=0}^{n+1} C_{n+1}^i a^{n+1-i} b^i$ доказано. Шаг индукции осуществлён, утверждение теоремы доказано.

Замечание. Выкладки в алгебраическом способе доказательства могут вам показаться непонятными и чрезмерно сложными. Но это не так. Выкладки здесь достаточно просты, но записаны на непривычном для вас языке, через символ суммирования и использование индексов. Вы можете сами переписать цепочку равенств из доказательства в привычных для вас выражениях, только тогда у вас между каждыми знаками равенства будут возникать многоточия. А чтобы лучше понять суть дела, рассмотрите переход для какого-нибудь маленького конкретного числа n например, $n = 2$.

Следующая теорема является следствием теоремы о биноме Ньютона. Её первый пункт получается из этой теоремы заменой b на $-b$. Остальные пункты — частные, хорошо известные вам со школы случаи, соответствующие значениям $n = 2$ и $n = 3$. Таким образом, доказывать эту теорему незачем.

Теорема 21. Верны следующие равенства (a и b — произвольные числа, $n \in \mathbf{N}$).

$$\text{а) } (a - b)^n = a^n - \frac{n}{1}a^{n-1}b + \frac{n(n-1)}{1 \cdot 2}a^{n-2}b^2 - \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3}a^{n-3}b^3 + \dots + (-1)^nb^n;$$

$$\text{б) } (a \pm b)^2 = a^2 \pm 2ab + b^2;$$

$$\text{в) } (a \pm b)^3 = a^3 \pm 3a^2b + 3ab^2 \pm b^3.$$

В качестве применения теоремы 20 докажем следующий факт.

Пример 31. Докажите, что для любого натурального числа n знакопеременная сумма $C_n^0 - C_n^1 + C_n^2 - \dots + (-1)^nC_n^n = 0$.

Доказательство. Применим теорему 20 к числам $a = 1$ и $b = -1$. Получим

$$(1 - 1)^n = \sum_{i=0}^n C_n^i \cdot 1^{n-i} \cdot (-1)^i = \sum_{i=0}^n (-1)^i C_n^i = C_n^0 - C_n^1 + C_n^2 - \dots + (-1)^n C_n^n,$$

откуда следует требуемый факт.

Упражнение 45. а) Докажите, что для любого натурального числа n верно равенство $C_n^0 + 2C_n^1 + 4C_n^2 + \dots + 2^n C_n^n = 3^n$;

б) Вычислите для любого натурального числа n знакопеременную сумму $C_n^0 - 2C_n^1 + 4C_n^2 + \dots + (-1)^n 2^n C_n^n$.

Говоря о биномиальных коэффициентах, невозможно не упомянуть красивую конструкцию, называемую *треугольником Паскаля*. Она строится следующим образом. Прежде всего формально введём число C_0^0 , которое положим равным единице, и запишем его в нулевой строчке. В первой строчке запишем последовательно числа C_1^0 и C_1^1 так, чтобы первое из этих чисел стояло чуть левее числа C_0^0 , второе — чуть правее. Во вторую строчку запишем числа C_2^0 , C_2^1 и C_2^2 в указанном порядке так, чтобы число C_2^0 располагалось левее числа C_1^0 , число C_2^2 — правее числа C_1^1 , а число C_2^1 — под числом C_0^0 . Аналогично заполним третью, четвёртую строки. И так далее, конца процессу нет. В результате получим бесконечный треугольник из чисел. Это и есть треугольник Паскаля. На рис. 1 изображены первые 10 строк (включая нулевую) этого треугольника (слева записаны биномиальные коэффициенты в стандартной форме, справа — их числовые значения); для наглядности элементы треугольника Паскаля помещены в шестиугольные клетки. Обычно этих клеток не рисуют, самих биномиальных коэффициентов не пишут, только их значения (как на правом рисунке). Строки чисел треугольника так и называют строками, но кроме этого в треугольнике выделяют столбцы, это ряды чисел, параллельные левой боковой стороне, ряды бесконечные. Нумеруются они также начиная с нулевого. Обратим внимание на то, что у чисел C_n^m , расположенных в одной строке, одинаков нижний индекс, и этот индекс совпадает с номером строки. Аналогично, у биномиальных коэффициентов, расположенных в одном столбце, один и тот же верхний индекс, равный номеру столбца. Иными словами, число C_n^m расположено на пересечении n -ой строки и m -го столбца.

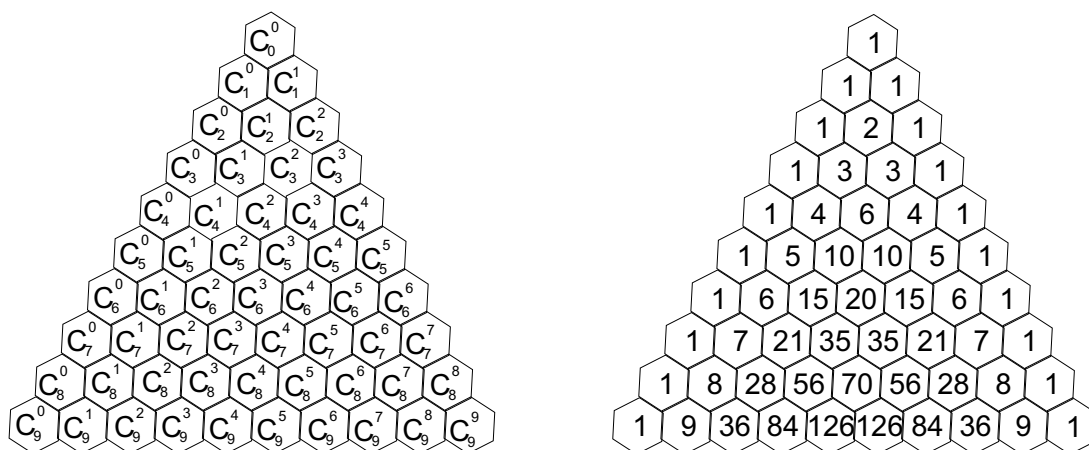


Рис. 1. Треугольник Паскаля

Чем же интересен такой треугольник, зачем он нужен. Давайте посмотрим, как свойства чисел C_n^m находят своё отражение в этом треугольнике. Прежде всего бросается в глаза симметричность треугольника — это следствие утверждения 3 из предыдущего пункта. Далее, по краям треугольника идут сплошь одни единицы, что тоже не удивительно, надо посмотреть на утверждение 1, чтобы понять в чём тут дело. Утверждение 2 говорит о том, что первый столбец треугольника представляет собой натуральный ряд. Более внимательные из вас могут заметить, что вторая строка треугольника тоже замечательна: она состоит из так называемых *треугольных* чисел (и все их содержит), т. е. из чисел, равных $1 + 2 + 3 + \dots + k$, $k \in \mathbb{N}$. Эти числа получили своё название из-за того, что они могут быть изображены геометрически следующим образом: возьмём правильный треугольник, разделим каждую из его сторон на n равных частей и через точки деления проведём прямые, параллельные сторонам треугольника — получим треугольную сетку. Оставим на рисунке только узлы этой сетки — их количество будет выражаться $n + 1$ -ым треугольным числом. То есть всякое треугольное число — это число точек в некотором построенном треугольнике.

Упражнение 46. Докажите, что для второй столбец треугольника Паскаля состоит из треугольных чисел и все их содержит.

Однако главная прелесть треугольника Паскаля следует из утверждения 4 предыдущего пункта. Согласно этому утверждению каждый элемент треугольника Паскаля (кроме самого верхнего) получается как сумма двух элементов, стоящих над ним в предыдущей строке. Это означает, что для заполнения следующей строки нет необходимости вычислять биномиальные коэффициенты длительным путём, через подсчёт факториалов, произведений. Достаточно просто осуществить несколько раз операцию сложения. Даже вручную не вызывает никакой сложности построение треугольника строки до пятнадцатой, а современный компьютер за доли секунды построит треугольник Паскаля хоть до миллионной строки, найдя тем самым *все* биномиальные коэффициенты (с нижним коэффициентом меньше миллиона) оптом. Вспомните, совсем недавно я вам обещал, что покажу способ быстрого нахождения чисел C_n^m . Этот способ перед вами.

Ну и наконец посмотрим на связь треугольника Паскаля с биномом Ньютона. Думаю, большинство из вас уже сообразило, что в n -ой строке треугольника Паскаля стоят коэффициенты в разложении суммы $(a + b)^n$. Таким образом, теперь вы можете легко раскрыть скобки в этом произведении. Например, чтобы раскрыть скобки в выражении $(x + 3)^5$ надо рассмотреть пятую строку треугольника Паскаля и получить

$$\begin{aligned}(x + 3)^5 &= x^5 + 5 \cdot x^4 \cdot 3 + 10 \cdot x^3 \cdot 3^2 + 10 \cdot x^2 \cdot 3^3 + 5 \cdot x \cdot 3^4 + 3^5 = \\ &= x^5 + 15x^4 + 90x^3 + 270x^2 + 405x + 243.\end{aligned}$$

Может случиться и так, что бином Ньютона поможет установлению свойств треугольника Паскаля. Например, как следствие примера 31, мы получаем, что знакопередающаяся сумма элементов каждой строки треугольника равна нулю.

Упражнение 47. Обратите внимание, что в треугольнике Паскаля есть строки, состоящие исключительно из нечётных чисел, например, строка 1, 7, 21, 35, 35, 21, 7, 1. Докажите, что таких строк бесконечно много, и укажите номера их всех.

Упражнение 48. Докажите, что всякое натуральное число, не равное единице, в треугольнике Паскаля встречается конечное число раз.

4. Сочетания с повторениями

Мы рассмотрели почти все основные понятия комбинаторики (сочетания, перестановки, размещения с повторениями и без), вывели формулы для их числа, привели примеры, кроме сочетаний с повторениями. Настал момент изучить и это понятие.

Как уже говорилось, весьма трудно дать теоретико-множественную трактовку сочетаниям с повторениями, а вот на языке выбора определение выглядит вполне естественно. Чтобы было проще понять формулу для числа \overline{C}_n^m , разберём сперва пример.

Пример 32. На автомобильном рынке продаются автомобили 5 фирм: «Ауди», «Тойота», «Мерседес», «Форд» и «Хонда» (автомобилей каждого типа неограниченное количество). Фирма желает приобрести 7 автомобилей. Сколькими способами может быть сделана такая покупка?

Решение. Нам требуется из пяти фирм — производителей автомобилей выбрать 7, причём порядок выбора не важен, а повторения допускаются. Значит, мы имеем дело с размещением с повторениями, и количество способов осуществить покупку равно \overline{C}_5^7 . Подсчитаем это число по-другому. Для этого зададимся вопросом, как по возможности более экономно (т. е. затратив возможно меньшее количество различных символов) записать какие машины и в каком количестве мы желаем приобрести. Пусть мы решили купить 3 «Тойоты», 2 «Мерседеса» и 2 «Хонды», а остальные машины решили не брать. Можно записать такую покупку, например, так: 3Т, 2М, 2Ф — разных символов 5. А можно так: ТТТММХХ — разных символов 3, меньше. Впрочем, если бы мы решили брать все типы автомобилей, то и при такой записи потребуется 5 символов. Можно ли меньше?

Можно. Давайте каждый автомобиль отмечать одним и тем же значком, например, крестиком. И договоримся, что первая серия крестиков соответствует «Ауди», вторая — «Тойоте» и т. д. Тогда указанная выше покупка

будет записана как $XXX0XX0XX$ (группы символов, соответствующие разным типам машин, должны отделяться друг от друга каким-либо символом, например, нулём, как в сделанной записи, или пробелом). Хорошо получится? Нет, не очень. Из нашей записи никак не увидать, какие два типа машин мы решили не брать. Выход прост. Надо ставить разделитель (сиречь ноль) после каждой группы машин, в том числе и после пустой. Получим последовательность $0XXX0XX00XX0$. Вот теперь порядок: первая группа пуста, т. е. «Ауди» мы не берём совсем, во второй — 3 символа, т. е. мы покупаем три «Тойоты», далее следуют 2 «Мерседеса», 0 «Фордов» и 2 «Хонды». Мы записали покупку, затратив всего два разных символа: крестик и нолик.

Теперь поймём, что общего во всех последовательностях крестиков и ноликов, которыми будут записываться допустимые покупки. Прежде всего заметим, что число крестиков равно числу покупаемых машин, а число ноликов — числу предлагаемых автомобилей. Значит, в последовательности будет 7 крестиков и 5 ноликов. Кроме того, последним символом обязательно будет нолик — символ, заканчивающий последнюю группу крестиков.

Сколько же есть таких последовательностей? У нас будет всего $5 + 7 = 12$ символов, последний — обязательно нолик. Среди остальных 11 крестиков ровно 7, и они могут стоять где угодно. Значит, их расставить можно C_{11}^7 способами. После этого положение ноликов (а значит, и вся последовательность) определяется однозначно.

Ответ: $\overline{C}_5^7 = C_{11}^7$ способов.

Теперь разберём общий случай.

Теорема 22. Для любых чисел $n \in \mathbf{N}$ и $m \in \mathbf{N} \cup \{0\}$ верно равенство $\overline{C}_n^m = C_{m+n-1}^m = C_{m+n-1}^{n-1}$.

Доказательство. Рассмотрим некоторое n -элементное множество $T = \{t_1, t_2, \dots, t_n\}$. Для каждого сочетания с повторениями построим последовательность из крестиков и ноликов следующим образом: сначала

столько крестиков, сколько раз в этом сочетании встретился элемент t_1 , потом один нолик, затем столько крестиков, сколько раз встретится элемент t_2 , затем снова нолик и т. д. до тех пор, пока не поставим n -ю группу крестиков, в которой крестиков столько, сколько раз встретился в сочетании элемент t_n . (Любая из групп крестиков, в том числе и последняя, может оказаться пустой.) Получим последовательность из m крестиков и $n - 1$ нолика. Построена функция из множества всех сочетаний с повторениями n -элементного множества по m в множество последовательностей крестиков и ноликов длины $m + n - 1$, содержащих ровно m крестиков. Эта функция, как легко проверить, является биекцией. Согласно принципу соответствия, получаем, что \overline{C}_n^m равно количеству таких последовательностей. Но каждая последовательность определяется набором мест, на которых стоят крестики (или нолики). Выбрать из $m + n - 1$ места m , которые будут занимать крестики (или, что то же самое, выбрать $n - 1$ место для ноликов) можно $C_{m+n-1}^m = C_{m+n-1}^{n-1}$ числом способов. Теорема доказана.

Упражнение 49. Номер автобусного билета шестизначный, отличный от 000000. Назовём билет *правильным*, если цифры его номера идут в неубывающем порядке. (Так, билеты с номерами 112569 и 333333 — правильные, а билеты с номерами 965211 и 921561 — нет.) Сколько существует правильных автобусных билетов?

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Барвайс Дж. Справочная книга по математической логике / Дж. Барвайс. В 4-х ч. Ч. 2. М. : Наука, Главная редакция физико-математической литературы. 1982.
2. Ткачук В. В. Математика — абитуриенту / В. В. Ткачук. 14-е изд., испр. и доп. М. : МЦНМО, 2007.
3. Виленкин Н. Я. Комбинаторика / Н. Я. Виленкин. М. : Наука. Главная редакция физико-математической литературы, 1969.
4. Виленкин Н. Я. «Рассказы о множествах» / Н. Я. Виленкин. 3-е изд. М. : МЦНМО, 2005.
5. Задачи по математике. Уравнения и неравенства / В. В. Вавилов, И. И. Мельников, С. Н. Олехник, П. И. Пасиченко. М. : Наука, 1987.

Содержание

ЛЕКЦИЯ I. Основы теории множеств	3
1. Понятие множества	3
2. Операции над множествами	6
3. Алгебра множеств	10
4. Декартово произведение множеств	13
ЛЕКЦИЯ II. Функции и графики	15
1. Основные определения	15
2. Действия над функциями. Простейшие числовые функции и их графики	20
3. Преобразования графиков функций	22
ЛЕКЦИЯ III. Обратная функция и её график	28
1. Композиция функций	28
2. Понятие функции, обратной к данной.	29
3. Инъекция, сюръекция и биекция. Связь биекции и обратной функции	31
ЛЕКЦИЯ IV. Тригонометрия	35
1. Единичная окружность и основные тригонометрические функции	35
2. Простейшие свойства основных тригонометрических функций	41
3. Тригонометрические формулы более сложного вида	45
4. Функции, обратные основным тригонометрическим, и их графики	52

ЛЕКЦИЯ V. Равносильные переходы в уравнениях (неравенствах)	57
1. Функциональный подход	
к уравнениям и неравенствам	57
2. Равносильные и неравносильные	
уравнения (неравенства)	59
3. Равносильные переходы в типичных	
уравнениях и неравенствах	65
4. Переходы к следствию. Равносильность	
и область допустимых значений	70
ЛЕКЦИЯ VI. Основы комбинаторики. Основные принципы	76
1. Конечные и бесконечные множества	76
2. Основные принципы комбинаторики	79
3. Математическая индукция	87
ЛЕКЦИЯ VII. Основы комбинаторики. Определения и формулы	94
1. Размещения, перестановки и сочетания	94
2. Биномиальные коэффициенты и их свойства	104
3. Бином Ньютона и треугольник Паскаля	107
4. Сочетания с повторениями	114
БИБЛИОГРАФИЧЕСКИЙ СПИСОК	117

Учебное пособие

МАТЕМАТИКА ДЛЯ ЭКОНОМИСТОВ

Составитель **Нохрин** Сергей Эрнестович

Редактор *О. В. Гусева*

Компьютерный набор *С. Э. Нохрина*

Компьютерная верстка *Е. В. Суховой*

Подписано в печать 14.09.2014. Формат 60 × 90 1/8.

Бумага писчая. Плоская печать. Усл. печ. л. 15,0.

Уч.-изд.л. 6,5. Тираж 50 экз. Заказ № 1553.

Издательство Уральского университета
Редакционно-издательский отдел ИПЦ УрФУ
620049, Екатеринбург, ул. С. Ковалевской, 5
Тел. 8 (343) 375-48-75, 375-46-85, 374-19-41
E-mail: rio@urfu.ru

Отпечатано в Издательско-полиграфическом центре УрФУ
620075, Екатеринбург, ул. Тургенева, 4
Тел. 8 (343) 350-56-64, 350-90-13
Факс: 8 (343) 358-93-06
E-mail: press-urfu@mail.ru

ОБ АВТОРЕ



НОХРИН
Сергей Эрнестович

Кандидат физико-математических наук, доцент кафедры моделирования управляемых систем Уральского федерального университета. Преподавательский стаж – 24 года (СИНХ, УЭМИИТ, УрФУ).